

34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
153 AATCTACCAAGTACAGGCAGACAAACCTATCTACAACTGTGGCTGAG 202
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
::: |||||
203 AAGCCCAAGAATATCAAGAAACACAGATATAAGGATATTTGGCCCTATGA 252
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
253 TTATAGCGGGTGAAGACTATCCCTGATAACCTCTGATGAGGATTCAGCT 302
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
303 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCRAGGCTTATATT 352
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
353 GCCACCACAGGCTCTTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 402
117 eTTPGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
403 TTGGGAATATAGTCTCTTATCATCTGTTATGGCATGCATGGAGTATGAAA 452
134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
453 TGGGAAGAAAAGTGTGACGCTACTGGCTGAGCCAGCAGAGATGCAG 502
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysse 167
|||||
503 CTGGAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAAGGAAATC 552
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
553 TGATATATATAATCAGGACTCTAAAGTTAAGTTCAATAGTGAACATCGAA 602
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
603 CTATCTACCAAGTTTCATTACAAGATTGGCCAGACCATGATGTACCTTCA 652
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCystyrGlnGl 217
|||||
653 TCTATAGACCCCTATTCTGTAGCTCATCTGGGATGTACGTTGTACCAGA 702
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
703 GGATGACAGTGTCCCATATGCATTTCACTGCAGTGTGGCTGTGGGAAGA 752
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
753 CTGGTGTATTTCGCTATT...GATTATACATGGATGTTGCTAAAGAT 799
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
800 GGGATAATTCCTGAGAACTTTCAGTGTGTTTTCAGTTGATCCGGGAAATGCG 849
267 qThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValT 284
|||||
850 GACACAGAGGCCCTTCATTTAGTTCAACGCGAGGAAACAATATGAACCTGGTCT 899
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
900 ACAATGCTGTATTAGAACTATTAAAGACAGACAGATGGATGTTATCAGAGAT 949
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
950 AAACATTTCTGGAAACAGAGATCAAGCAAGCAATTTGTTATTCCTCGAGAAAA 999
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
1000 TCACACTCTCCAGCAGACACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1049
334 hrLysAlaLysMetMetAsnGlnArgThrLysMetGluIleLys 350
|||||

1050 CAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAATGGAAATCAAA 1099
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1100 GAATCTTCTTCTCTTACTTTAGGACTTCTGAAATTAAGTGCAAAAAGAA 1149
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1150 GCTAGTTTTGCACCCCTGCTAAATCAAGCAGCTTCTTTTGACTTTCTGGAGC 1199
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1200 TAAATACAGTTTTGACAAAAATGCTGACACAACCATGAATGGAATGGCAGACA 1249
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1250 AAGGCATTTTCCAATAGTTGGGAGCCTCTTCAGAAGCATCAAAAGTTTGA 1299
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1300 TTTGGGCTCTCTTTTGTGTTGAGGATGTTCTTAATTTCTAAACCTGTAATG 1349
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1350 CAGCAGGAAGATATTTAATTCAAAGGTGCAATAACACGGACCAAAATCA 1399
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1400 ACTCTCTTTGAATTTGATACAGCAGAGAGAAACCAAGAGGTGGACAGCAA 1449
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1450 GGAACACTTTTCTTATTTGGAATCTCAACCATCATGATCTTGTTTTGTAG 1499
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1500 AGATGCAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATAT 1549
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1550 TCATCTGCCATATGACTCTAAACACCAAAATACGTAATGCCCTCTAATGTAAA 1599
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1600 GCACCATGACTCTAGTCTCTTGGTGTATATTCTTACATACCTTTAGTGG 1649
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
|||||
1650 AAAATCCTTATTTTTCATCATGCTCCCAAGTGTACCAGTTCTTAAGATG 1699
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1700 TCTCTTGATTTACTCTGAGAAGAGATGGAACTGTTTTTTCCTTCTCTCT 1749
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1750 GTTGCCACATCTCTACATCCCTCTCTCTTCTTATTACAAATTCACATGATT 1799
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1800 CTTTATCACTGAATTTCTCCAACCAATATTTCTCTCACTATTATGAACCCAGGAG 1849
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1850 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATGAATCCCCCTCC 1899
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1900 ACTTCTGTACGACACACTGAATCATTTATTGTTGGTTGAGGAAGCTGGAG 1949
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||

1950 AATTCTCACCRAAATGTTCCCAAACTCTTATCTCAGCTGTGAAGGTAAAA 1999
651 ileGlyThrSerLeuGluTrpGlyThrSerGluProLysLysPheAs 667
|||||
2000 ATTGAACATCACTGGAATGGGTGGAACATCTGAACCAAGAAATTGGA 2049
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2050 TGACTCTGTGATCTAGTACCAAGCAAGAGTGTAAAACTCCGAAGTCTTA 2099
684 ySer 685
|||||
2100 AATCA 2104

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-951-260A-1

seq_documentation_block:
; Sequence 1, Application US/08951260A
; Patent No. 6004791
; GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; TITLE OF INVENTION: AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,260A
; FILING DATE: October 16, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No 6004791ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELE: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to tRNA
US-08-951-260A-1

alignment_scores:
Quality: 748.00 Length: 357
Ratio: 3.041 Gaps: 7
Percent Similarity: 68.908 Percent Identity: 43.137

alignment_block:
US-09-600-358A-4 x US-08-951-260A-1 ..

Align seg 1/1 to: US-08-951-260A-1 from: 1 to: 2226
3 GlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLysLysI 19
|||||
37 CAATCGGACCTAGTAGGAGGCTTCTTGGAGCAGCAGGAGGCCCGGACCA 86
19 eThrLysGluGlu...PheAlaAsnGluPheLeuLysLeuLysArgGlns 35
|||||
87 CCGGAAGGGGCAATCTCTCGCCGCTGACTTCAGCGACATTAAAGGCCGCT 136
35 eThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGluAsn 51
|||||
137 CAGTGGCTTGAAGACTCAAGGTGTGTCTCCACTAAAGCGCGGAGTCCAG 186
52 AlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAspTy 68
|||||
187 CAGGAAACTCAAGAGAACCGCTACAAAGACGGTGTACCGTATGATGA 236
68 rSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSertYrI 85
|||||
237 GACGAGAGTCATCTTCCCTGCTCCAGGAGGAGGACACGAGATTACA 286
85 leAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIleAla 101
|||||
287 TTAATGCCAATTTCATCGGGGCACAGATGGAAGCCAGGCTACATTGCG 336
102 ThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTr 118
|||||
337 ACAGAAGGACCCCTGCTCACACTCTGTGTGACTTCTGGCGCCTGGTTTG 386
118 pGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluMetG 135
|||||
387 GGAGTTGGAAATCAAGGTGATCTTGATGCGCTGTCCAGGAGACAGAAATG 436
135 lLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu 151
|||||
437 GACGAGGAAGTGTCAAGCTACTTGGGCCACAGAGCGGGAG...CCTCTA 483
152 GluPheGlyProPheSerValSerCysGluAlaGluLysArg...LysSe 167
|||
484 CAGGCCGGGCGCTTCTGCATCACCCTGACAAAGGAGACAGCAGTACTTC 533
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
534 GCATCACTCTCAGGACCTCCAGGTTCATTCCAGAGGAATCCCGTC 583
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
584 CTGTGCACACAGTACAGTACATGTCTTGGCCGACCCAGCGGTTCCAGC 633
201 SerIleAspProIleLeuGluIleTrpAspValArgCysTyrGlnG 217
|||
634 AGTCCGATCACATCTCACCATGTGGAGGAGGCCCTTGCCTCCAAAG 683
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
684 ACTTGGACCTGGACCCCTCTGTGCCACTGCAGTCTGGCTGTGGACGAA 733
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
734 CAGGTGTCTTGTGTCT...GTGATTACGTGAGGCAGTGTCTTCTGACT 780
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
781 CAGACAAATCCCAACCAATTCAGCCCTCTTGAAGTGGTCTCTGGAGATCG 830
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||
831 GAACAGCGACCTGCGCGGTGTCAGACAGACAGGACGATACAGGTTCTCTGT 880
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp..... 296
|||
|||

```
881 ACCACAGTGGTCACTATTCTCCGCACTCTCCAGAACAAACAGTCCC 930
297 .....ValileArgAspLysH1 302
931 CTCTACCAGAACCTCAAGGAGAACCGGCTCCATCTGCAAGGACTCCTC 980
302 sSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnH1st 319
111 .....CTGCTGCACATCCGCC 1021
981 GTCCCTCAGGACCTCCCTCAGCC.....CTGCTGCACATCCGCC 1021
319 hrLeuGlnAlaAspSerTyrSer.....ProAsnLeu 329
111 ..... 1071
1022 CACTGGGTGGCTTCTCAGGAGCATCTCGGTGCTGGGCCACCGACCTT 1071
330 ProLysSerThrThrLysAla 336
1072 CCCATGGCTGACACTTACGCT 1092
```

seq_name: /cqn2_6/ptodata/1/ina/6B_COMB.seq:US-08-821-278A-1

```
seq_documentation_block:
; Sequence 1, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P10101
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-821-278A-1
```

```
alignment_scores:
Quality: 736.50 Length: 327
Ratio: 3.069 Gaps: 6
Percent Similarity: 73.394 Percent Identity: 44.037
```

alignment_block:
US-09-600-358A-4 x US-08-821-278A-1 ..

```
Align seg 1/1 to: US-08-821-278A-1 from: 1 to: 1529
5 GluIleLeuGlnLysPheLeuAspGluAlaGlnSerLysLysIleThrLy 21
111 ..... 34
34 GACTTGGTGAGGAGCTTCTGGAGCACTGGAGCCCGCGGACTACCGGA 83
21 sGluGlu...PheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrL 37
111 ..... 84
84 GGGGGCAATCTTCGTGCTGAGTTCAGCGACATTAAGGCCCGCTCAGTGG 133
37 ysTyrLysAlaAspLysThrThrProThrValAlaGluAsnAlaLys 53
111 ..... 134
134 CCTGGAAGTCTGAAGGTGTGTTCACATAAGCCGCGCATCGCGTGGG 183
54 AsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerAr 70
111 ..... 184
184 AACAGCAACAGAACCGCTACAAAGATGTGTAGCATATGATGAGACAAG 233
70 gValGluLeuSerLeuIleThrSerAspGluAspSerTyrIleAsnA 87
111 ..... 234
234 AGTCATCCTTCTCCCTGCTCCAAAGAGGAGGACATGAAATATACATCAATG 283
87 laAsnPheIleLysGlyValTyrGlyProLysAlaTyrIleAlaThrGln 103
111 ..... 284
284 CCAACTTCATCCGGGGGATGAGTACGACCGCCCTACATGTCGACGCAA 333
104 GlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTy 120
```

```
111 ..... 334
334 GGACCCCTGCCTCACACACTGTTGGACTTCTGGCGCTGTTGGGAGTT 383
120 rSerValLeuIleValMetAlaCysMetGluTyrGluMetGlyLysL 137
111 ..... 384
384 TGGGGTCAAGGTAATCTGATGCCCTGTCGAAGAGACAGAAAATGGACGGA 433
137 ysLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPhe 153
111 ..... 434
434 GGAAGTGTGAACCTATTTGGCCCCGGAGCAGGAG...CCTCTAAAGGCT 480
154 GlyProPheSerValSerCysGluAlaGluLysArg...LysSerAspTy 169
111 ..... 481
481 GGGCCTTCTGCATCACCTCGACAAAGGAGACAACTGAATGCAGACAT 530
169 rIleIleArgThrLeuLysValLysPheAsnSerGluThrArgThrIle 186
111 ..... 531
531 CACTCTCAGGACCTCCAGGTTACATTTCCAGAAGGAATTCGCTGTGTC 580
186 yrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerIle 202
111 ..... 581
581 ACCAACTACAGTATATGCTCTGCCAGACCCCGGGTTCACAGCAGTCT 630
203 AspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGluAspAs 219
111 ..... 631
631 GATCACATCTCACCATGTTGGAGGAGGCCGCTGCTCCCAAGGGCTTGG 680
219 pSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGly 236
111 ..... 681
681 ACCTGGACCCCTGTTGTCACATGTCAGTGTGCTGCGGACGACACAGGTG 730
236 alIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyIle 252
111 ..... 731
731 TCCTGTGCGCT...GTGACTATGTGAGGAGTGTGCTGCTGACCCACACA 777
253 IleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThrGl 269
111 ..... 778
778 ATCCCTCCCAACTTCAGTCTCTTCCAAAGTGTCTCTGGAGATGCGGAAACA 827
269 nArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrAsnA 286
111 ..... 828
828 GCGGCTGCGAGCTGCAGACAGAGGAGGAGTACAGGTCTCTGTACCACA 877
286 laValLeuGluLeuPheLysArgGlnMetAspValIleArgAspLysHis 302
111 ..... 878
878 CAGTGGCTCAGCTATCTCCGCGC.....ACTCTCCAGGACACACAGC 918
303 SerGlyThrGluSerGlnAlaLysHisCysIlePro.....GluLysAs 317
111 ..... 919
919 CCCCAATACCAAGAACCTCAAGGAGAACTGCGCTCCAATCTGCAAGGAAGC 968
317 nHisThrLeuGlnAlaAspSerTyrSerPro 327
111 ..... 969
969 TTTTCTCCCTCAGGACCTCTCCAGCCCTCAGCCCTGCGCT 999
seq_name: /cqn2_6/ptodata/1/ina/6A_COMB.seq:US-08-854-585-1
seq_documentation_block:
; Sequence 1, Application US/08854585
; Patent No. 6114140
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```

3795 GTGAGAGATTGG.....CCTCAAGCAGGCTCAGGACTATGAGAC 3838
156 PheSerValSerCysGluAlaGluLysArgLysSerAspTyrIleIleAr 172
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3839 ATAACGTGTGCAATGACATCAGAAATGTTCCTCCGGAATGGACCATCAG 3888
172 gThrLeuLysValLys.....PheAsnSerGluThrArgThrIleTyrG 187
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3889 AGATTTCACAGTCAAAATATCCAGACAAAGTAGAGATCAACCTCTGAGAC 3938
187 InPheHisTyrLysAsnTrrProAspHisAspValProSerSerIleAsp 203
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3939 AGTTCATTTTACCCTCCTGGCAGACACCGGTGTCCTCCGACACCATGAC 3988
204 ProIleLeuGluLeuIleTrrAspValArgCysTyr.....GlnGluAs 218
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
3989 CTGCTCATCACTCCGGTACCTCGTTCGTGACTACATGAAGCAGAGTCC 4038
218 pAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrG 235
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
4039 TCCCGAATCCGCGATTCTGGTCATTGCAATGCTGCTGGGTGCGGAAGACGG 4088
235 IyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGly 251
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
4089 GCACITTCATTGCCATTGATCGTCTCATCTACCAGATAGAGAATGAGAAC 4138
252 IleIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgTh 268
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
4139 ACCGTG.....GATGTGATGGATTGTGTATGACCTTCGAAT 4176
268 rGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrA 285
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
4177 GCATAGGCCCTTAATGGTGCAGACAGACGACCATGATGTTTCTCCTCAATC 4226
285 snAlaValLeuGluLeuPheLysArgGlnMetAsp 296
4227 AGTGTGTTTGGATATTGTGCAGATCCAGAAAGAC 4261
seq_name: /cgn2_6/ptodaat/1/ina/PCTUS_COMB.seq: PCT-US95-05512-1
seq_documentation_block:
; Sequence 1, Application PC/TUS9505512
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine
; TITLE OF INVENTION: Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05512
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5117 base pairs

```


US-09-600-358A-4 x US-08-018-129-4

Align seg 1/1 to: US-08-018-129-4 from: 1 to: 2790

```
14 AlaGlnSerLysLysIleThrLysGluLupheAlaAsnGluPheLeu 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
836 GCTGAGACCAAGTAAAGTCAAAACAGCGCTTTGGGAGAAATTTGAGAC 887
30 sLeuLysArgGlnSerThrLysLysAlaAspLysThrTyrProThr 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
888 ACTACAAACAGAGAGTCAAACTT.....CTCTACAGCCGAA 925
47 hrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAsp 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
936 AAGAGGGTCAAAAGCAAGCAAAACAAAAATAGATATAAAACATC 975
64 LeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
976 CTGCCCTTTGATCATACCAAGGGTTGCTCAGCATCTGTGATCCCAATGA 1025
80 uAsp....SerSerTyrIleAsnAlaAsnPheIleLysGlyValTyr.... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1026 GCCTGTTTCAGATACATCAATGCAATATCATCATGCTCAATTTGAAA 1075
95 .....GlyProLysAlaTyrIleAlaThrGlnGly 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1076 CCAAGTGCAACAATTCAAAGCCCAAAAGAGTTACATTTGCCACAGGC 1125
105 ProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSe 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1126 TGCCTGCAAAACAGGTGAATGACTTTTGGCGGATGGTGTCCCAAGAAA 1175
121 rValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLys 138
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1176 CTCGGGTGATGTGTAGCAAGAAAGAGTGGAGAGAGAAAGATA 1225
138 yscysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu...GluPhe 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1226 AATGTGCTCAATACTGG.....CCTGATGAGTATGCTCTAAAGAAAT 1269
154 GlyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrIl 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1270 GCGCTCATGCTGTGTAGCAAGCTCAAAAGAAAGCGCGCTCATGACTATAC 1319
170 eIleArgThrLeuLysVal.....LysPheAsnSerGluThrA 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1320 CTAAGAGAACTTAACCTTCAAGAGTTGGACAGGAATACGGAG...A 1366
183 rgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPro 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1367 GAACGGTCTGGCAATACCACCTTCGGACCTGGCGGACCACCGCGTGCC 1416
200 SerSerIleAspProIleLeuLeuLeuIleTrpAspValArgCysTyrGl 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1417 AGCGACCTGGGGCGCTGCTGACTTCTGGAGGAGGTGCACCAATAGCA 1466
216 nGluAsp.....AspSerValProIleCysIleHisCysSerAlaGlyC 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1467 GGAGAGCATCATGATGACGGCGCGCTGCTGGTGCACGTGCAGTCTGGAA 1516
231 ysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeu 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1517 TTGGCCGACAGGACGCTC...ATTGATGATGATATCTTATTTCACATC 1563
248 LeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIleAr 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1564 ATCAGAGAGAAAGGTGTGTGACTGCGATATTGACGTTCACAAACCATCCA 1613
264 gGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrG 281
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1614 GATGTGCGGTCTCAGAGGTGAGGATGGTCCAGACAGAACACAGTACC 1663
281 luLeuValTyrAsnAlaVal.....LeuGluLeuPheLysArgGln 294
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1664 GATTATCTATATGCGCGTCCAGCATTTATTGAAACACTACAGCGCAGG 1713
295 MetAsp..... 296
1714 ATTGAAGAAGAGACAGAAAAGCAAGAGGAAGGCCACGAATATACAAATAT 1763
297 .....ValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysH 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1764 TAAGTATTCTCTAGCGGACAGACGAGTGGAGATCAGAGCCCTCTCCCGC 1813
311 isCysIleProGlu..... 315
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1814 CTGTGACTCTCTGCGCACCTGTGCAGAAATGAGAGAAAGAGTTCAGATGAGA 1863
316 .....LysAsnHisThrLeuGlnAlaAspSerTyrSer.....Pr 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1864 GTCATGAAAACGTGGCGCTGTGATGCAACAGCAGAAAAGTTTCAGATGAGA 1913
327 oAsnLeuProLys.....SerThrThrLysAlaA 337
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1914 AAACCTGCCAAAACCTTCAGCAGACAAATAGATGTGGACTTTCACCTCTCC 1963
337 laLysMetMetAsnGlnGlnArgThrLysMetGluIleLysGluSerSer 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1964 CTAAAAGATCAGGACGACGCAAGAAAGTTTATGTGAAGTCTGAATTT 2013
354 SerPheAsp.....PheArgThrSerGluIleSerAlaLysGl 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2014 GGATTTGGAAGGCTTGCAATGTGGTTGACTACCTTTTGATAAGCAAAAT 2063
366 uGluLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuG 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2064 TGAACCACTTAAAGACCACTGATTTTAACCTCAACAATACCTGCTTCCC 2113
383 luLeuAsnTyrSerPheAspLys 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2114 AATTACTCATTTCTCCACATAAG 2136
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-448-250-4
seq_documentation_block:
; Sequence 4, Application US/08448250
; Patent No. 5981251
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP ID: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,250
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/018,129
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-017
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 130..1911
; US-08-448-250-4

alignment_scores:
  Quality: 496.00      Length: 441
  Ratio: 1.968        Gaps: 15
  Percent Similarity: 57.143      Percent Identity: 31.293

alignment_block:
US-09-600-358A-4 x US-08-448-250-4 ..
Align seg 1/1 to: US-08-448-250-4 from: 1 to: 2790

14 AlaGlnSerLysLysLysLysLysGluGluPheAlaAsnGluPheLeuLy 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
838 GCTGAGACCACAGATAAAGTCAAACAGGCTTTGGGAAGAAATTTGAGAC 887
30 sLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
888 ACTACAAACACAGAGTGCAACTT.....CTCTACAGCGCAA 925
47 hrValAlaGluAsnAlaLysAsnLysLysLysAsnArgTyrLysAspIle 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
926 AAGAGGGTCAAGGCAAGCAAGAAACAAAAACAAATAAGATATAAAACATC 975
64 LeuProTyrAspTyrSerArgValGluLeuSerLeuLysLysLysLys 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
976 CTGCGCTTTTCATCATACCGAGGTGTCTCCTCAGCATCTGTGATCCCAATGA 1025
80 uAsp...SerSerTyrIleAsnAlaAsnPheIleLysGlyValTyr.... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1026 GCCTGTTTCAGATTACATCAATGCAATATATCATCGCTGGAATTTGAAA 1075
95 .....GlyProLysAlaTyrIleAlaThrGlnGly 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1076 CCAAGTGCAACAATTCAAAGCCCAAAAGAGTTACATTGCCACACAAGGC 1125
105 ProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSe 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1126 TGCCTGCAAAACACGGTGAATGACTTTTGGCGGATGGTGTCCAAGAAA 1175
121 rValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLys 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1176 CTCCCGAGTGATGTGATGACAGCAAGAAAGTGGAGAGGAAAGAGTA 1225
138 ysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu...GluPhe 153
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1226 AATGTGCTCAATACTGG.....CCTGATGAGTATGCTCTAAAGAAATAT 1269
154 GlyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrII 170
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1270 GCGGTGATGCTGTGTAGAACGTCAAAGAAAGCGCCGCTCATGACTATAC 1319
170 eileArgThrLeuLysVal.....LysPheAsnSerGluThrA 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1320 GCTAAGAGAACTTAACCTTCAAGGTTGGACAAGGGAATACGGAG...A 1366
183 rgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPro 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1367 GAACGGCTCTGGCAANTACCACTTTTCGGACCTGGCCGGACCAACGCGGTGCC 1416
200 SerSerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGI 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1417 AGCGACCTGGGGCGTGTGGACTTCCTGGAGGAGTGCACCATTAAGCA 1466
216 nGluAsp.....AspSerValProIleCysIleHisCysSerAlaGlyC 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1467 GGAGAGCATCATGATGCAGGGCGCGTGTGGTGGCAGCTGCAGTCTGGAA 1516
231 ysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeu 247
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1517 TTGGCCGGACAGGGACGTTTC...ATTGTGATTGATATCTTATTATGACAT 1563
248 LeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIleAr 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1564 ATCAGAGAGAAAGGTGTGACTCGGATATTGACGTTCCCAAAACCATCCA 1613
264 gGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluLysTrp 281
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1614 GATGTTGCGGTCTCTCAGAGGTCAGGATGGTCCAGACAGAAAGCACAGTACC 1663
281 luleuValTyrAsnAlaVal.....LeuGluLeuPheLysArgGln 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1664 GATTTATCTATATGCGCGTCCAGCATTTATATTGAAACACTACAGCGCAGG 1713
295 MetAsp..... 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1714 ATTGAAGAGAGCAGAAAGCAAGACAGAAAGGCGCAGCAATATACAATAT 1763
297 .....ValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysH 311
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1764 TAAGTATTTCTCTAGCGGACCAAGAGTGGAGATCAGAGCCCTCTCCCGC 1813
311 iScysIleProGlu..... 315
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1814 CTTGTACTCTCTCCGCCACCCCTGTCAGAAATAGAGAAAGACAGTGTAGA 1863
316 .....LysAsnHisThrLeuGlnAlaAspSerTyrSer.....Pr 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1864 GTCTATGAAAGAGTGGCGCTGATGCACACAGCAGCAAGAAAGTTTCAGATCAGA 1913
327 oAsnLeuProLys.....SerThrThrLysAlaAla 337
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1914 AAACCTGCCAAAACCTTCAGCACAGAAATAGATGTGGACTTTTCACCTCTCC 1963
337 lalysMetMetAsnGlnArgThrLysMetGluIleLysGluSerSer 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1964 CTAAGAGATCAGGACCAAGCAAGAAAGTTTATGTGAAGTCTGAATTT 2013
354 SerPheAsp.....PheArgThrSerGluIleSerAlaLysGI 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2014 GGATTTGGAAGCGTTGCAATGTGTGCTGACTTACCTTTTGTGATAAGCAAAATT 2063
366 uGluLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuG 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2064 TGAACCATTTTAAAGACCACTGTATTTTAACCAACAATACCTGCTTCCC 2113
383 luleuAsnTyrSerPheAspLys 390
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2114 AATTACTCATTTCTCAGATAAG 2136
seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-449-644-4
seq_documentation_block:
; Sequence 4, Application US/08449644
; Patent No. 5856162
; GENERAL INFORMATION:
; APPLICANT: Schlessler, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam

```

```

alignment_scores:
  Quality: 492.50      Length: 467
  Ratio: 1.954         Gaps: 16
  Percent Similarity: 53.961  Percent Identity: 29.122

alignment_block:
  US-09-600-358A-4 x US-08-449-644-4  ..

  Align seg 1/1 to: US-08-449-644-4 from: 1 to: 4651

21  LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
    |||||::: :: ||| ::| |||||:::
2662 AAAGAGGAATATGACAGCTTTT .....GAAGGACAGTCACGATC 2702
    |||||::: ::| ||| |||||::: ::|

37  sTyrLysAlaAspLysThrTyProThrValAlaGluAsnAlaLysA 54
    ::| |||||::: ::| |||||::: ::|

2703 TTGG.....GATCTAGCTAAAAAGATCAAA 2728

54  snlleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArg 70
    || |||||::: ::| ||| |||||::: |||||::: |||||:::
2729 ATAGACCAAAACCGATATGGAACATTTATAGCATATGATCATTCCAGA 2778
    |||||::: ::| ||| |||||::: |||||::: |||||::: |||||:::

71  ValGluLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAl 87
    || |||||::: ::| ||| |||||::: |||||::: |||||::: |||||:::
2779 GTGATTTTGCAACCGTAGAGGATGATCCTCTCAGATTATATTATGTC 2828
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::

87  aasnPheileLysGlyValTyrGlyProLysAlaTyrIleAlaThrGlnG 104
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::

2829 CAATATTATGATGCTACCGACAGACCAAGTCATTACATTGCAACCCAAG 2878

```

104 lyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleLysGluTyr 120
||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
GTCCGGTTTCATGAACAGCTATGATTCTTGAGGATGATTTTGCAGAAG 292879

121 SerValLeuIleIleValMetAlaCysMetGluTyrGlnMetGlyLysLy 137
::: ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
CAAPCTCGTTGCAPTGTGATGGTTTACAATAATTAAGTTAGAGTGGCCGGGT 2978

137 sLYScysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPheG 154
||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
TAAATGCTATAAATATTGGCTGTGATGATCATGAAGTT.....TAGT 3019

154 lyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrIle 170
||| : ||||| : ||||| : ||||| : ||||| : ||||| :
GTGACTTCAAAGTAACGTGTGTAGAATAATGCAACCATTGCTGAATATGTA 3069

171 Ilearg.....ThrLeuIlysValLysPheAsnSereGluThrArgHrII 185
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
GTTAGGACATTCACCCPGGAAGAGGGGGGTACAATAAGTAATCGCTGAAGT 3119

185 eTYrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerI 202
||| : ||||| : ||||| : ||||| : ||||| : ||||| :
TAAACAGTTCATTTACGGGCTGGCTGCACCATGGNGTGCCCTACCATG 3169

202 leaspProIleLeuGluLeuIleTPasPValArgCystYrGlnGluasp 218
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
CTACAGGGCTGCTTTCCTTTATCCGGCGAGTCAAGTTATCAAAACCCCTCCC 3219

219 AspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGI 235
::::: ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
AGTGTGCCGCCATCGTGTGACATTGCAGTCGTGGTCTGGAGCAACTGG 3269

235 yVallieCysAlaIleValAspTyrThrTpmetLeuLeuLysAspGLyI 252
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
CTGCTACATTGTGATTCACATCATCTAGACATGGCTGAAAGAGAGGGTG 3319

252 lelleProGluAsnPheSerValPheSerLeulleArgGluMetArgThr 268
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
TTGTT.....GATATTTACAATTGTCTCAAGAGCTTTAAGATCT 3357

269 GlnArgProSerLeuValGlnThrGlnGluTyrGluLeuValTyraS 285
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
CGCGGTATTAAATATGGTCCAAGAGAGAACAGTACATTTTTTATTATCATGA 3407

285 nAlaValLeuGluLeu..... : ||||| : ||||| : ||||| : ||||| : ||||| :
TGCCATTTAGACCGCTGTTATGTGGAGNAACGTCACATCTGCTGTG 3457

291 ..PhetyS...ArgGInMetAspValileArg..... : ||||| : ||||| : ||||| : ||||| : ||||| :
AATTTAAAGCTGCATATTTTGGATGATGATTAGATAGACTCCACAGACTAAC 3507

300 AsplyshisSerGlythrGlnSerGlnAlaLysHis..... : ||||| : ||||| : ||||| : ||||| : ||||| :
TCTTCACATCTCAAGGATGAATTCATTACACTCTCAANTCAGTCAACCCTCG 3557

312 : ||||| : ||||| : ||||| : ||||| : ||||| :
ACTACAAGCTGAAGACTGCAGTATACGGTGCCTGCCA..AGNAAACNTG 3604

319 hrLeuGln.....AlaAspSerTyrSerProasn.....LeuproLys 331
::: ||| : ||||| : ||||| : ||||| : ||||| : ||||| :
ACAAGAACCGTTTTTCATGGACATGCTGCCACCTCACAGATGCTCTGCCTTTT 3654

332 SerThrThrLysalaalalyssMetMetAsnGlnlArgThrLysMet... 347
||| : ||||| : ||||| : ||||| : ||||| : ||||| :
TTAATTACAATTGATGGGAGACGAATACATCAATGCTCTCTTAT 3704

348 ...GluiLeLysgluserSerSerPheaspPheArgThrSerGluiles 363
:::||| : ||||| : ||||| : ||||| : ||||| : ||||| :
GGACAGCTACAGCGACACCGACTCTTTC..... : ||||| : ||||| : ||||| : ||||| : ||||| :
erAlalyssGluGluLeuValLeuHisProAlalyssSerSerThrSerPhe 379

```

3733 .....ATCGTCACACAATACCCCTCTGCCAAACACTGTAAAA 3768
      ::::: ::::: :::::
380 AspPheLeuGluLeuAsnTyrSerPhe..... 388
      ::::: ::::: :::::
3769 GACTTCTGGAGATTAGTGTATGATGCTGCTACCTCCATGTCATGTT 3818
      ::::: ::::: :::::
389 .....AspLysAsnAlaAspThrThrMetLysTrpGlnThrLysA 402
      ::::: ::::: :::::
3819 AAACGAAGTCGACTGTCTCCAGGGCTGCCCTCAGTACTGGCCAGAGGAG 3868
      ::::: ::::: :::::
402 laPhe.....ProfileValGlyGluProLeuGlnLysHisGln 414
      ::::: ::::: :::::
3869 GGATGCTACGATATGGCCCAACCAAGTGAATGATGTCTGTGTTCAATG 3918
      ::::: ::::: :::::
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
      ::::: ::::: :::::
3919 GACTGTGATGTGATCAACGGGATTTTATAGGATATGCAATCTAACAAAGACC 3968
      ::::: ::::: :::::
431 O 431
3969 A 3969

seq_name: /cgn2_6/ptodata/1/1na/5B_COMB.seq:US-08-087-244A-4
seq_documentation_block:
; Sequence 4, Application US/08087244A
; Patent No. 5863755
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Sap, Jan M.
; APPLICANT: Ullrich, Axel
; APPLICANT: Vogel, Wolfgang
; APPLICANT: Fuchs, Miriam
; APPLICANT: Fuchs, Miriam
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-KAPPA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,244A
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4317
US-08-087-244A-4
```

```

alignment_scores:
  Quality: 492.50      Length: 467
  Ratio: 1.954        Gaps: 16
  Percent Similarity: 53.961  Percent Identity: 29.122

alignment_block:
US-09-600-358A-4 x US-08-087-244A-4 ..
Align seg 1/1 to: US-08-087-244A-4 from: 1 to: 4651

21  LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
      |||||::: ::::: |||
2662 AAAGAGGAATATGAGAGCTTTT.....GAAGGACAGTCAGCATC 2702
      ::::: ::::: :::::

37  sTyrLysAlaAspLysThrTyrProThrThrValAlaGluAsnAlaLysA 54
      ::::: ::::: :::::
2703 TTGG.....GATGTAGCTAAAAAAGATCAAA 2728
      ::::: ::::: :::::

54  snileLysLysAsnArgTyrLysAspilleLeuProTyrAspTyrSerArg 70
      |||||::: ::::: |||
2729 ATAGAGCAAAACCGATATGAAACATATATAGCATATGATCCTCCAGA 2778
      ::::: ::::: :::::

71  ValGluLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAl 87
      |||||::: ::::: |||
2779 GTGATTTTGCACACCGTAGAGGATGCTCTCCTCAGATTATATTAATGC 2828
      ::::: ::::: :::::

87  aasnPheileLysGlyValTyrGlyProLysAlaTyrIleAlaThrGlnG 104
      |||||::: ::::: |||
2829 CAACATATATTGATGGCTACCAGAGACCAAGTCATTACATTTGCAACCCAAG 2878
      ::::: ::::: :::::

104  LyProLeuSerThrThrLeuLeuAspPheTrpArgMetileTrpGluTyr 120
      |||||::: ::::: |||
2879 GTCCCGTTCATGAACAGTGATGATTCTGGAGGATGATTTGGCAAGAA 2928
      ::::: ::::: :::::

121  SerValLeuIleileValMetAlaCysMetGluTyrGluMetGlyLysLy 137
      ::::: ::::: :::::
2929 CAATCTGCTTGCTGTGTGTGTACAAATTTAGTTGAGTTGGCGGGT 2978
      ::::: ::::: :::::

137  sLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeuGluPheG 154
      |||||::: ::::: |||
2979 TAAATGCTATAAATATTGGCCTGATGATCAAGTT.....TATG 3019
      ::::: ::::: :::::

154  LyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrile 170
      |||||::: ::::: |||
3020 GTGACTTCAAAGTAAAGTGTAGAAATGGAACCACTTCTGCAATATGTA 3069
      ::::: ::::: :::::

171  IleArg.....ThrLeuLysValLysPheAsnSerGluThrArgThrI 185
      ::::: ::::: :::::
3070 GTTAGGACATTCACCTCGAAAGAGGGGTACAATGAAATCCGTGAAGT 3119
      ::::: ::::: :::::

185  eTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSerSerI 202
      |||||::: ::::: |||
3120 TAAACAGTTCCATTTACCGGCTGGCTGCATGAGTGAGTCCCTACCATG 3169
      ::::: ::::: :::::

202  leAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp 218
      ::::: ::::: :::::
3170 CTACAGGGCTGCTTTCTTTATCCGGCGAGTCAAGTTATCAAAACCTCCC 3219
      ::::: ::::: :::::

219  AspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGl 235
      ::::: ::::: :::::
3220 AGTGTGTGCCCATCGTTGATATTGCTGCTGCTGCTGCTGCTGCTGCTG 3269
      ::::: ::::: :::::

235  yValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyI 252
      |||||::: ::::: |||
3270 CTGCTACATTTGATTTGATGATCATGCTAGCATGCTGAAAGAGAGGGTG 3319
      ::::: ::::: :::::

252  leIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThr 268
      ::::: ::::: :::::
3320 TTGTT.....GATATTACAATTTGTCAAGCCCTTAAGATCT 3357
      ::::: ::::: :::::

269  GlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValTyrAs 285
      ::::: ::::: :::::
```



```

109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1135 ACAGTCAATGACTTCTGGCAGATGGCGTGGCAGGAGAACAGCGGTGCAT 1184
||||| ||||||| ||||||| ||||||| ||||||| |||||||
125 eValMetAlaCysMetGluTyrGluMetGlyLysCysGluArgT 142
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1185 CGTCATGACCAACCCGAGAGGTGGAGAAAGCGCGACAAATCGTCCCAT 1234
||||| ||||||| ||||||| ||||||| ||||||| |||||||
142 yTrpAlaGluProGlyGluMetGlnLeuGluPheGlyProPheSerVal 158
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1235 ATGGCCCCAGGTGGC...ATGCAGCTGCTTATGGGCCCTACTCTGTG 1281
||||| ||||||| ||||||| ||||||| ||||||| |||||||
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1282 ACCAAGCTCGGGGAGCATCACACACCGAATACAACTCCGTACCTTACA 1331
||||| ||||||| ||||||| ||||||| ||||||| |||||||
175 sVallys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1332 GTCTCTCCCGCTGGACAATGGAGACCTGATTCGGGAGATCTGGCATTACC 1381
||||| ||||||| ||||||| ||||||| ||||||| |||||||
189 isTyrLysAsnTrpProAspHisAspValProSerSerIleAspProIle 205
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1382 AGTACCTGAGCTGGCCCGACCAATGGGTGCCAGTGAGCTGGGGGTGTC 1431
||||| ||||||| ||||||| ||||||| ||||||| |||||||
206 LeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp.....AspSe 220
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1432 CTCAGCTTCTGGACCACTCAACCAAGCGGCAAGAAAGTGCCTCACGC 1481
||||| ||||||| ||||||| ||||||| ||||||| |||||||
220 rValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyValI 237
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1482 AGGGCCCATCATCTGTCACCTGCGCGCGCGCATCGCGCCGACAGGCACCA 1531
||||| ||||||| ||||||| ||||||| ||||||| |||||||
237 leCysAlaIleValAspTyrThrTrpMetLeuLeuLysAspGlyIleIle 253
|| ||||| ||||||| ||||||| ||||||| |||||||
1532 TCATTGTCTATC.....GACATGCTC 1551
||||| ||||||| ||||||| ||||||| ||||||| |||||||
254 ProGluAsnPheSerValPheSerLeu..... 262
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1552 ATGGAGAACATCTCCACCAAGGCGCTGGACTGTGACATTGACATCCAGAA 1601
||||| ||||||| ||||||| ||||||| ||||||| |||||||
263 ...IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnG 278
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1602 GACCATCCAGATGGTGGCGGCGCGCGCTCGGGCATGGTGCAGACGGAGG 1651
||||| ||||||| ||||||| ||||||| ||||||| |||||||
278 LuGlnTyrGluLeuValTyrAsnAlaVal.....LeuGluLeuPhe 291
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1652 CGCAGTACAGTTCATCTACGTGGCCATCGCCCGCTCAATGAAACCACT 1701
||||| ||||||| ||||||| ||||||| ||||||| |||||||
292 LysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluSerGl 308
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1702 AAGAAGAAGCTGGAGTCTCTGCAGTCCGAG...AAGGGCCAGAGTCGGA 1748
||||| ||||||| ||||||| ||||||| ||||||| |||||||
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSerT 325
|| ||||| ||||||| ||||||| ||||||| |||||||
1749 GTACGGGAACATC.....ACCT 1765
||||| ||||||| ||||||| ||||||| ||||||| |||||||
325 yTrSerProAsnLeuProLysSerThrThrLysAlaAlaLys..... 338
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1766 ATCCCCACCCATGAAGATGCGCCATGCCAAGGCCCTCCCGCCTCGTCC 1815
||||| ||||||| ||||||| ||||||| ||||||| |||||||
339 .....MetMetAsnGlnGlnArgThrLy 346
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1816 AAACACAAGGAGGATGTGTATGAGAACCTGCACACTAAGAACAAGAGGA 1865
||||| ||||||| ||||||| ||||||| ||||||| |||||||
346 smetGluIleLysGluSerSerPheAspPheArgThrSerGluIles 363
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1866 GGAGAAGATGAAGAACCGCTGCACAGACAAAGGAGAGCAAGGGTT 1915
||||| ||||||| ||||||| ||||||| ||||||| |||||||
363 erAlaLysGluGlu 367
|| ||||| ||||||| ||||||| ||||||| |||||||
1916 CCTCAAGAGGAAG 1929
||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

seq_name: /cgn2_6/ptodata/1/1na/6A_COMB.seq:US-09-358-685-1

```

seq_documentation_block:
; Sequence 1, Application US/09358685
; Patent No. 6121047
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF SHP-1 EXPRESSION
; FILE REFERENCE: RTS-0081
; CURRENT APPLICATION NUMBER: US/09/358,685
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (256)..(2049)
US-09-358-685-1

```

```

alignment_scores:
  Quality: 485.00      Length: 388
  Ratio: 2.012        Gaps: 14
  Percent Similarity: 62.113  Percent Identity: 32.990

```

```

alignment_block:
US-09-600-358A-4 x US-09-358-685-1 ..

```

Align seg 1/1 to: US-09-358-685-1 from: 1 to: 2277

```

15 GlnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLe 31
||||| ||||||| ||||||| ||||||| ||||||| |||||||
964 GAGTCGGAGGATACAGCAAGCGTGGCTTCGGGAGGATTTGAGAGTTT 1013
||||| ||||||| ||||||| ||||||| ||||||| |||||||
31 uLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThrThrv 48
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1014 GCAGAAGCAGGAGGTCAAG.....AACTTGACACGAGCGTCTGG 1051
||||| ||||||| ||||||| ||||||| ||||||| |||||||
48 alAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeu 64
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1052 AAGGCGAGCGGCCAGAGAAAGAGGCGCAAGACCGCTACAAAGAACATTC 1101
||||| ||||||| ||||||| ||||||| ||||||| |||||||
65 ProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp...Gl 80
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1102 CCCTTGACACAGCGGAGTGATCTCTGCAGGACGGGACAGTAACATCCC 1151
||||| ||||||| ||||||| ||||||| ||||||| |||||||
80 uAspSerSerTyrIleAsnAlaAsnPheIleLysGly...ValTyrGlyP 96
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1152 CGGGTCCGACTACATCAATGCCAATCATCAAGAACACGCTGCTAGGCC 1201
||||| ||||||| ||||||| ||||||| ||||||| |||||||
96 ro.....LysAlaTyrIleAlaThrGlnGluProLeuSerThr 108
|| ||||| ||||||| ||||||| ||||||| |||||||
1202 CTGATGAGAACGCTAAGACCTACATCGCAGCGAGGCTGTCTGGAGGCC 1251
||||| ||||||| ||||||| ||||||| ||||||| |||||||
109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1252 ACGGTCAATGACTTCTGGCAGATGGCGTGGCAGGAGAACAGCGGTGCAT 1301
||||| ||||||| ||||||| ||||||| ||||||| |||||||
125 eValMetAlaCysMetGluTyrGluMetGlyLysLysCysGluArgT 142
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1302 CGTCATGACCAACCCGAGAGGTGGAGAAAGCGCGAACAATCGTCCCAT 1351
||||| ||||||| ||||||| ||||||| ||||||| |||||||
142 yTrpAlaGluProGlyGluMetGlnLeuGluPheGlyProPheSerVal 158
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1352 ACTGGCCCCGAGGTGGGC...ATGCAGCGTGTATGGGCCCTACTCTGTG 1398
||||| ||||||| ||||||| ||||||| ||||||| |||||||
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
||||| ||||||| ||||||| ||||||| ||||||| |||||||
1399 ACCAACTGCGGGAGCATGACACACCCGAATACAAACTCCGTACCTTACA 1448
||||| ||||||| ||||||| ||||||| ||||||| |||||||
175 sVallys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
||||| ||||||| ||||||| ||||||| ||||||| |||||||

```



```
1449 GGTCTCCCGCTGACAAATGGAGACCTGATTCCGGGAGATCTGGCATTACC 1498
189 iStyLysAsnTrpProAspHisAspValProSerSerIleAspProIle 205
1499 AGTACCTAGAGTCCCGCAGCAGTGGGTGCCCTCCCTAGCTGGGGGTGTC 1548
206 LeuGluLeuIleTrpAspValArgCysTrpGlnGluAsp.....AspSe 220
1549 CTCAGCTTCTGGACACAGATCAACACCGCGCAGGAAGTCTGCCTCAGCG 1598
220 rValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyValI 237
1599 AGGCCCATCATCTGCTGACCTGCAGCGCGGCATCGCGCCGACAGGCACCA 1648
237 leCysAlaIleValAspTrpThrTrpMetLeuLeuLysAspGlyIleIle 253
1649 TCATTGTGCATC.....GACATGCTC 1668
254 ProGluAsnPheSerValPheSerLeu..... 262
1669 ATGGAGAACATCTCCACCAAGGCGCTGGACTGTGCACATTGACATCCAGAA 1718
263 ...IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnG 278
1719 GACCATCCAGATGTGGCGGCGCAGCGCTCGGCATGTGTGCAGACGGAGG 1768
278 LuGlnTrpGluLeuValValTrpAsnAlaVal.....LeuGluLeuPhe 291
1769 CGCAGTACAAGTTCTACTGCTGCCATCGCCCTCCCTAGTTTCATTGAAACCACT 1818
292 LysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluSerGl 308
1819 AAGAAGAGTGGAGGTCTCTGCAGTCCAG...AAGGCCAGGAGTCCGA 1865
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSer 325
1866 GTACGGGAACATC.....ACCT 1882
325 yrSerProAsnLeuProLysSerThrThrLysAlaAlaLys..... 338
1883 ATCCCCAGCCATGAAGAAATGCCATGCCATCGCCCTCCCGCACCTCGTCC 1932
339 .....MetMetAsnGlnGlnArgThrLy 346
1933 AAACACAAAGGAGTGTGTATGAGAACCTGCACACTAAGAACAAAGAGGGA 1982
346 sMetGluIleLysGluSerSerPheAspPheArgThrSerGluIleS 363
1983 GGAGAAAGTGAAGACGAGCGGTCCAGCAGACAGGAGGAGCAAGGAGGTT 2032
363 erAlaLysGluGlu 367
2033 CCCTCAAGAGGAG 2046
```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:us-08-202-389-7

seq_documentation_block:

Sequence 7, Application US/08202389

Patent No. 5536636

GENERAL INFORMATION:

APPLICANT: Freeman Jr., Robert M.

APPLICANT: Plutsky, Jorge

APPLICANT: Neel, Benjamin G.

APPLICANT: Rosenberg, Robert D.

TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

```
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,389
FILING DATE: 28-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,926
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,141
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/721,112
FILING DATE: 26-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: BIH92-05MA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 145..2037
US-08-202-389-7
```

alignment_scores:

Quality: 484.00 Length: 443

Ratio: 1.847 Gaps: 18

Percent Similarity: 59.142 Percent Identity: 31.603

alignment_block:

US-09-600-358A-4 x US-08-202-389-7 ..

Align seg 1/1 to: US-08-202-389-7 from: 1 to: 2143

```
15 GlnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLe 31
:||||: :||| :||| :||| :||| :||| :||| :||| :|||
847 GAGTCCGAGGATACAGCAAGGCTGGCTTCTGGGAGAGTTTGAGAGTTT 896
31 uLysArgGlnSerThrLysLysAlaAspLysThrThrProThrThrV 48
:||||: :||| :||| :||| :||| :||| :||| :||| :|||
897 GCAGAAGCAGGAGGTGAG.....AAGTGCACGAGCGTCTGG 934
48 alAlaGluAsnAlaLysAsnIleLysLysAsnArgTrpLysAspIleLeu 64
:||||: :||| :||| :||| :||| :||| :||| :||| :|||
935 AAGGGCAAGCGCCAGAGAACAAGGCAAGAACCGCTACAAGAATCTCTC 984
65 ProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAsp...G1 80
||||:||||:||||:||||:||||:||||:||||:||||:||||:
985 CCCTTTGACCACAGCCGAGTGATCTCGAGGCGGACAGTAAACATCCC 1034
80 uAspSerSerTrpIleAlaAsnPheIleLysGly...ValTyrGlyP 96
||||:||||:||||:||||:||||:||||:||||:||||:||||:
1035 CGGTCCGACTACATCAATGCCAACTACATCAAGAACCAGCTGTAGGCC 1084
96 ro.....LysAlaTyrIleAlaThrGlnGlyProLeuSerThr 108
|||:||||:||||:||||:||||:||||:||||:||||:||||:
1085 CTGATGAGAAGCGCTAAGACCTACATCCAGCCAGGCGTGTCTGGAGGCC 1134
```

109 ThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle1 125
1135 ACAGTCAATGACTTCGGCAGATGGCGTGGCAGGAGAACAGCGGTGCAT 1184
125 eValMetAlaCysMetGluTyrGluMetGlyLysLysCysGluArgT 142
1185 CGTCATGACCCAGCGAGGTGGAGAAAGCGGAAACAAATCGTCCCAT 1234
142 yTrpAlaGluProGlyGluMetGlnLeuGluPheGlyProPheSerVal 158
1235 ACTGGCCCGAGGTGGC...ATGCAGCGTGTATGGCCCTACTCTGTG 1281
159 SerCysGluAlaGluLysArgLysSerAspTyrIleIleArgThrLeuLy 175
1282 ACCAACTGGGGAGCATCACACACCGCAATACAACTCCGTACCTTACA 1331
175 sValLys.....PheAsnSerGlu...ThrArgThrIleTyrGlnPheH 189
1332 GGTCCTCCCGCTGGACAATGGAGACCTGATTCGGGAGATCTGGCAATTACC 1381
189 iStryLysAsnTrpProAspHisAspValProSerSerIleAspProIle 205
1382 AGTACCTGAGCTGGCCGACCATGGGGTCCCGAGTGGCGTGGGGGTGC 1431
206 LeuGluLeuIleTrpAspValArgCysTyrGlnGluAsp.....AspSe 220
1432 CTCAGCTTCCTGGACCATCAACACCGCGCAGGAAAGTCTGCCTCACGC 1481
220 rValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyValI 237
1482 AGGGCCCATCATCGTCACTGCAGCGCGCGCATCGCGCACAGGCACCA 1531
237 leCysAlaIleValAspTyrThrTrpMetLeuLeuLysaspGlyIleIle 253
1532 TCATTGTCATC.....GACATGCTC 1551
254 ProGluAsnPheSerValPheSerLeu..... 262
1552 ATGGAGAACATCTCCACCAAGGCGCTGGAGTGTGACATTGACATCCAGAA 1601
263IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnG 278
1602 GACCATCCAGATGGTGGCGCGCAGCGCTGGCATGGTGCAGACGGAGG 1651
278 luGlnTyrGluLeuValTyrAsnAlaVal.....LeuGluLeuPhe 291
1652 CGCAGTACAGTTCATCTACGTGGCCATCGCCCGCATTCATTGAACCACT 1701
292 LysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluSerGl 308
1702 AAGAAGAAGCTGGAGTCTGTCAGTGCAG...AAGGCCAGAGTCCGA 1748
308 nAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAlaAspSerT 325
1749 GTACGGGAACATC.....ACCT 1765
325 yrSerProAsnLeuProLysSerThrThrLysAlaAlaLysMetMetAsn 341
1766 ATCCCCGAGCCATGAGATGATCCCATGCCAAGGCCTCC..... 1803
342 GlnGlnArgThrLysMetGluIleLysGluSerSerSerPheAspPheAr 358
1804CGCACCTCGTCCAAACACACAGGAGGATGTGTATGAGAACCTGCA 1847
358 gThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAlaLys 375
1848 CACTAAG.....AACAGAGGAGGAA..... 1869
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
1870AGTGAAGAA 1878

392 AlaAspThrThrMetLysTrpGlnThrLysAlaPheProIleValGlyGl 408
1879 GCAGCGGTGAGCAGACAGGAGAGCAAGAGGTTCCCT..... 1917
408 uProLeuGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluG 425
1918CAAGAGGAGTGAAGC.....GGTGTGTCTCTCAGGTGGC 1951
425 lYcysSerAsnSerLysProValAsnAla 434
1952 CATGCTCAGCCCTGACCTGTGGAAGCA 1980
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-202-389-11
seq_documentation_block:
; Sequence 11, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: Freeman Jr., Robert M.
; APPLICANT: Plutsky, Jorge
; APPLICANT: Neel, Benjamin G.
; APPLICANT: Rosenberg, Robert D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: B1H92-05MA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..1893
US-08-202-389-11

alignment_scores:
Quality: 483.00 Length: 488
Ratio: 1.851 Gaps: 20


```

COUNTRY: U.S.A.
ZIP: 10036
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/449,644
  FILING DATE: 24-MAY-1995
  CLASSIFICATION: 514
  PRIORITY APPLICATION DATA:
    APPLICATION NUMBER: US 08/087,244
    FILING DATE: 01-JUL-1993
  ATTORNEY/AGENT INFORMATION:
    NAME: Coruzzi, Laura A.
    REGISTRATION NUMBER: 30,742
    REFERENCE/DOCKET NUMBER: 7683-042
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 212-790-9090
    TELEFAX: 212-869-8864/9741
    TELEX: 66141 PENNIE
  INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 4374 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: unknown
    MOLECULE TYPE: DNA (genomic)
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1..4371
    US-08-449-644-3

alignment_scores:
  Quality: 478.50      Length: 473
  Ratio: 1.899        Gaps: 17
  Percent Similarity: 53.277      Percent Identity: 28.541

alignment_block:
  US-09-600-358A-4 x US-08-449-644-3 ..
  Align seg 1/1 to: US-08-449-644-3 from: 1 to: 4374

21  LysGluGluPheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrLy 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2698 AAAGAGAAATACGAGAGCTCTTT.....GAAGGCCAGTCAGCCTC 2738

37  sTyrLysAlaAspLysThrTyrProThrValAlaGluAsnAlaLysA 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2739 TTGG.....GATGTGGCTAAAGAGATCAAA 2764

54  snlleLysLysAsnArgTyrLysAspIleLeuProTyrAspTyrSerArg 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2765 ACAGAGCAAGAACCGATACGGAACATTTATCGCATATGATCACTCCAGA 2814

71  ValGluLeuSerLeuIleThrSerAspGluAspSerSerTyrIleAsnAl 87
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2815 GTCATCTGCAACCTGTGGAAGATGACCTCTTCAGATTACATTATATGC 2864

87  aAsnPhelle.....LysGlyValTyrGlyProLysA 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2865 CAACATACATGACATTGGCTGTACAGGATGGCTACCGACCAAGGCC 2914

98  latyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrp 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2915 ACTACATTGCAACTCAAGCCCGAGTTTCATGAACCGTATATGATTTTGG 2964

115 ArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysMetGl 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2965 AGGATGTGTGGCAAGACGAGCTGCTGCTGATTGTGATGTCCTACTAATTT 3014

```

```

131  uTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluProGlyG 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3015 AGTGGAAAGTGGCCGGTGAATGCTATAAATATTGGCCTGATGACTAG 3064

148  luMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGluLys 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3065 AGGTT.....TATGGTGAATTCAAAGTCACTCGCTAGAAATGGAG 3105

165  ArgLysSerAspTyrIleArg.....ThrLeuLysValLysPheAs 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3106 CCACCTTGCTGAGTATGTCGTTAGACATTCACCTTGGAAAGGAGGGCTA 3155

179  nSerGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspH 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3156 TAATGAATCCGTTGAAGTCAACAGTTCACCTTCACTGGCTGGCCTGACC 3205

196  isaspValProSerSerIleAspProIleLeuLeuIleTrpAspVal 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3206 ATGGTGTTCATACACCGAACAGAGGCTCTGTCTATTATCCGGAGAGTC 3255

213  ArgCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAl 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3256 AGCTATCTAACCCCTCCAGTGTGGGCCCATTTGTGTACACATGACAGTC 3305

229  aglyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpM 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3306 TGGTGTGGCGCACAGGCTGTACATTGTTATTGACATAATGCTGGACA 3355

246  etLeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3356 TGGCTGAAAGAGAGGGTGTGGTT.....GACATCTACAACCTGT 3393

263  IleArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGl 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3394 GTGAAAGCCCTTACGATCTCGGCCCATTAATATGTTAGTACAGAGGAACA 3443

279  nTyrGluLeuValTyrAsnAlaValLeuGluLeu..... 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3444 GTACATTTTATTTCATGATGCCATTTTAGAAGCCTGCTTATGTGGAGAAA 3493

291  .....PheLys....ArgGlnMetAspValIleArg 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3494 CTGCCATCCCTGTGTGTGAATTTAAAGCTGCATATTTTGTATGATTCGA 3543

300  .....AspLysHisSerGlyThrGluSerGlnAlaLys 310
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3544 ATGACTCTCAGACTAACTCTCTCATCTCAAAGATGAATTCAGACTCT 3593

310  sHis.....CysI 313
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3594 GAATTCGGTCAACCCCTCGACTCAAGCTGAAGACTGCAGCATAGCCTGCC 3643

313  leProGluLysAsnHisThrLeuGln.....AlaAspSerTyrSerPro 327
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3644 TGCCA...AGGAACCATGACAAGAACCGTTTTCATGATATGCTCCACCT 3690

328  Asn.....LeuProLysSerThrThrLysAlaAlaLysMetMetAsnGl 342
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3691 GACAGATGCTGCTCTTTTAAATTAATCAATTTGATGGGAGAGCAGTAACTA 3740

342  nGlnArgThrLysMet.....GluIleLysGluSerSerSerPheAsp 357
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3741 CATCAATGCTGCTCTTATGATAGCTATAGCAGCCAGCAGCTTC.... 3786

357  heArgThrSerGluIleSerAlaLysGluLeuValLeuHisProAla 373
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3787 .....ATCGTCAACAATACCCA 3804

374  LysSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPhe..... 388
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3805 CTGCCAAACACTGTGAAAGACTTCTGGAGATTAGTATATGATTCAGGATG 3854

389  .....AspLysAsnAlaAspThrThrM 396

```

Align seg 1/1 to: US-08-087-244A-3 from: 1 to: 4374

21	LYSGLUGLUPHEALAAASNGLUPHELEULYSLEULYSARGGLNPSERTHRLY	37
2698	AAGAGAAATACGAGAGCTCTTT.....GAAGGCCAGTCAGCCCTC	2738
37	STYRLYALAAASPLYSTHRTYRPROTHRVALLAAGLUASNALALYSA	54
2739	TTGG.....GATGTGGCTAAAGGATCAA	2764
54	SNLELYSLYASNAARGTYRLYLSASPLLEUPROTYRASPTYRSERARG	70
2765	ACAGAGCAAGAACCGATACGGAACATATTCGCATATGCATCACTCCAGA	2814

seq_documentation_block:

; GENERAL INFORMATION:

; APPLICANT: Schles:

APPLICANT: ullrich, Axel

APPLICANT: Fuchs, Miriam

NOVEL RECEIVED	TITLE OF INVENTION:	TITLE OF INVENTION:
PHOSPHATASE		

; NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS.

ADDRESSEE: PENNIE & EDMONDS
STREET, 1155 AVONING OF + 10 AM

CLTY: New York
STATE: New York

COUNTRY: U.S.A.

; COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

; SOFTWARE: PatentIn Release #1

APPLICATION NUMBER: US/08/087

CLASSIFICATION: 435

NAME: Coruzzi, Laura A.

REFERENCE/DOCKET NUMBER: 7683

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8804/5741
TELEX: 66141 PENNTE

```

; INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:

```

LENGTH: 43/4 base pairs
TYPE: nucleic acid

STRANDEDNESS: double

; MOLECULE TYPE: DNA (genomic)

; NAME/KEY: CDS

US-08-087-244A-3

Quality: 478.50

Percent Similarity: 53.277

alignment block:

03-03-000-358A-4 X 03-06-067-244A-3

```
3494 CTGCCATCCCTGTGTGTGAATTTAAAGCTGCATATTTTGTATATGATTCGA 3543
300 .....AspLysHisSerGlyThrGluSerGlnAlaLys 310
3544 ATAGACTCTCAGACTAACTCCTCTCACTCAAGAAGATGAATTCAGACTCT 3593
310 sHis.....CysI 313
   ::
3594 GAATTCGGTCACCCCTCGACTACAAGCTGAAGACTGCAGCATAGCCTGCC 3643
313 leProGluLysAsnHisThrLeuGln.....AlaAspSerTyrSerPro 327
3644 TGCCA...AGGAACCATGACAGAACCGTTTCATGGATATGCTCCACCT 3690
328 Asn.....LeuProLysSerThrThrLysAlaAlaLysMetMetAsnG1 342
   ::
3691 GACAGATGCTGCCTTTTAAATTAACAATTGATGGGGAGAGCAGTAACCTA 3740
342 nGlnArgThrLysMet.....GluIleLysGluSerSerSerPheAspP 357
   ::
3741 CATCAATGCTGCTCTTATGGATAGCTATAGGCAGCCAGCAGCTTC.... 3786
357 heArgThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAla 373
   ::
3787 .....ATCGTCACACAATACCCA 3804
374 LysSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPhe..... 388
   ::
3805 CTGCCAAACACTGTGAAGACTTCTGGAGATTAGTATATGATTACGGATG 3854
389 .....AspLysAsnAlaAspThrThrm 396
   ::
3855 TACCTCCATCGTGATGCTAAATGAAGTGGACCTGCTCAGGGCTGCCCCAC 3904
396 etLysTrpGlnThrLysAlaPhe.....ProIleValGlyGlu 408
   ::
3905 AGTACTGGCCAGAGAAGGAATGCTGCGATATGGTCTATCCCAAGTGGAA 3954
409 ProLeuGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluG1 425
   ::
3955 TGTATGCTTGTTCATGGACTGTGATGTGATCAATCGAATTTTATGAAT 4004
425 yCysSerAsnSerLysPro 431
   ::
4005 ATGCAACCTAACGAGACCA 4023
```

OM of: US-09-600-358a-4 to: N_Geneseq_032802.* out_format : pfs
Date: May 26, 2002 2:18 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cn2.1/USPTO.spool/US09600358/runat_24052002.101555.27056/app_query.fasta_1.761
-DB=N_Geneseq_032802 QPWT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -QCAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YCAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09600358 @CGN1_1_339 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-600-358a-4
Query length: 692
Database: N_Geneseq_032802.*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 351.650000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AA930696				3615.00	5772
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AA930696				3573.00	5701
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AAV81742				3543.50	5652
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH79342				2977.50	4748
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:AB199910				1104.50	1739
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:AAF21801				1058.00	1674
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:AAV17097				748.00	1170
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:AAV03112				736.50	1156
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:AAV17099				708.00	1103
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:AAV71076				509.00	785
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:AA706027				508.00	775
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAI58418				506.50	774
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:AA45353				497.00	765
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:AA592612				496.00	757
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AAZ08539				496.00	755
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:AAQ72913				492.50	752
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:AA735306				487.00	752
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH99616				486.00	747
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAK52292				486.00	740
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAK53276				486.00	740
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAK560778				486.00	739
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26454				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26470				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26471				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26484				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26485				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:AAK32917				485.00	750
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:AAA09637				485.00	748
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:AAZ32916				485.00	747
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:AAV71075				485.00	746
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:AA735307				483.00	745
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH73761				482.00	744
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH98348				482.00	734
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:AA591774				478.50	726
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:AA569143				478.00	736
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAH73240				478.00	722
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:AAQ94311				477.50	720
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26466				477.00	738
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26469				477.00	738
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26473				477.00	738

/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26475 + 477.00
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:AAV05762 + 476.50
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AAV81897 + 476.50
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:AAQ71842 + 476.00
/net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:AAF26467 + 475.00

seq_name: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AA930696

seq_documentation_block:

ID: AAX90696 standard; cDNA; 2356 BP.

XX AAX90696;

XX 01-OCT-1999 (first entry)

XX Human Cytoplasmic phosphatase Lyp2, cDNA.

XX Lymphoid Protein Tyrosine Phosphatase gene; Lyp gene; immunosuppressant;
KW intracellular tyrosine phosphatase; pPase; lymphocyte; fetal liver;
KW resting lymphoid cell; protein tyrosine kinase; PKs; transgenic animal;
KW lymphoma; T cell antigen receptor signalling; intronic sequence;
KW cytokine receptor signalling; protein tyrosine phosphatase; pPase; ds.
XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 42..2120

XX /*tag= a

XX /product= "Cytoplasmic phosphatase, Lyp2"

XX /note= "Expressed in fetal liver tissue and resting lymphoid cells"

XX WO936548-A1.

XX 22-JUL-1999.

XX 18-JAN-1999; 99WO-CA00038.

XX 16-JAN-1998; 98CA-2220853.

XX (HSCR-) HSC RES & DEV LP.

XX Roifman CM;

XX WPI: 1999-444404/37.

XX P-PSDB; AAY28653.

XX New nucleic acid encoding intracellular tyrosine phosphatase and related proteins, used to modulate signalling through T cells, particularly as immunosuppressant

XX Claim 5b; Page 54; 105pp; English.

XX The present cDNA sequence is that of the Lymphoid Protein Tyrosine Phosphatase gene, Lyp2 that encodes a cytoplasmic tyrosine phosphatase protein with a single catalytic domain. The Lyp gene has been localised to human chromosome 1p13. It is expressed significantly in fetal liver and in resting lymphoid cells. The Lyp2 gene sequence arises by alternative splicing of the Lyp1 mRNA and the coding sequence reads into the intronic sequence of Lyp1 till a termination codon is encountered. The Lyp gene products are important for regulation of T cell antigen and cytokine receptor signalling and for early and late stages of T cell differentiation. Lyp2 has immunosuppressive activity. This sequence is used to treat lymphoma, to reduce or prevent T cell activation or proliferation and to control thymocyte differentiation. Fragments of this gene are used, as probes or primers, to identify allelic variations and for diagnosis of diseases associated with reduced Lyp activity. Transgenic animals in which a Lyp gene has been inactivated or replaced by a heterologous Lyp gene are used to study and screen for possible therapeutic agents.

XX Sequence 2356 BP; 750 A; 466 C; 466 G; 674 T; 0 other;

SQ

alignment_scores:
Quality: 3615.00 Length: 692
Ratio: 5.224 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-600-358A-4 x AAX90696 ..

Align seg 1/1 to: AAX90696 from: 1 to: 2356

```
1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
42 ATGACCAAAAGAGAAATTCGAGAGAGTTCTCGATGAGGCCCAAGCAA 91
|||||
17 sLysIleThrLysGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92 GAAAATTTACTAAAGAGAGAGTTTCCCAATGAATTTCTGAAGCTGAAAGGC 141
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 AATCTACCAAGTACAGGCGACAGCAAAACCTTATCTCAACTGTGGCTGAG 191
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 AATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGCCCTATGA 241
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242 TTATAGCCGGGTAGAATATCCCTGTATAACCTCTGTATGAGGATTCACGCT 291
|||||
84 YrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCCAAGGCTTATATT 341
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
342 GCCACCACGGGCTCCTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 391
|||||
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTGCTCTTATCATTTGTTATGGCATGCATGGATATGAAA 441
|||||
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442 TGGAAAGAAAAGTGTGAGGCTACTGGCTGAGCCAGGAGATGCAG 491
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492 CTGGAATTTGGCCCTTTCTCTGTATCTGTGAAGCTGAAAAAGGAAATC 541
|||||
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TGATATATATACAGGACTCTAAAAGTTAAGTTCAATAGTGAACATCGAA 591
|||||
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTACCAAGTTTCATTACAAGAAATTGGCCAGACCATGATGTACCTTCA 641
|||||
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
642 TCTATAGACCCCTATTCTTGAGCTCATCTGGGATGTACGTTGTTACCAAGA 691
|||||
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGAGTTTCCCATATATGCATTTCACTGCAGTGGCTGGTGGGAAGA 741
|||||
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTGTGGCTATTGTTGATTATACATGGATGTTGCTAAAAGAT 791
|||||
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
```

```
792 GGGATAAATTCCTCAGAACTTCACTGTTTTTTCAGTTTGCATCGGAAATGCG 841
|||||
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCCCTTCATTAGTTCAAACGCGCAGGAACAATATGAACCTGTCT 891
|||||
284 YrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACAATGCTGTATTAGAACTATTAAAGACAGACAGATGGATGTTATCAGAGAT 941
|||||
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTTCTGGAACAGAGAGTCAAGCAAAAGCATTTGTATTCTCAGAAAAA 991
|||||
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAAGCAGACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAACGACCAAAAATGATGACCAACAAGGACAAAAATGGAATCAAA 1091
|||||
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1092 GAATCTCTCTCTTCTGACTTTAGGACTTCTGAAATAAGTGCAAAAGAAGA 1141
|||||
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTGCACCCCTGCTAAATCAAGCAGCTTCTTTGACTTTCTGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAAATTACAGTTTTGCACAAAAATGCTGACACAACCATGAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAATAGTTGGGAGCCCTCTTCAGAAAGCATCAAAAGTTTGA 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTTGTGGGGAGTGTCTTAATCTTAAACCCGTAAATG 1341
|||||
434 laLaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CACGAGGAAGATATTTTAATTCAAAGGTGCCAATAACACGACCAAAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCCTTTTGAATTTGATACAGACAGAGAAAAACCAAGAGGTGGACAGCAA 1441
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAATACTTTCTTATTTTGAATCTCAACCACATGATCTCTGTTGTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGACGGCTCAAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATTAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACCTGCCATATGACTCTAAACACCAAAATACGTAATGCCCTCTAATGTAAA 1591
|||||
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTGCTCTTGGTGTATATTCTTACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGCTCCCAAGTGGTACCAGTTCTAAGATG 1691
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTTACCTGAGAAGAGATGGAACGTGTTTTTCTCTCTCTCTCT 1741
|||||
```


567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSers 584
|||||
1742 GTTGCCACATCCTACATCCTCTCTCTATTATACATTCACATAGTT 1791

584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTCCTCAACCAATATTTCTCATTATGAACCCAGGAG 1841

601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCACTCTCCAAGGATAGATGATGAATCCCCCTCC 1891

617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1892 ACTTCTGTACGACACCTGAATCATTTATCTGCTGAGGAAGCTGGAG 1941

634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAATGTTCCCAATCCTTATCTCAGCTGTGAAGGTAAAA 1991

651 IleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGAACATCACTGGAATGGGTGGACATCTGAACCAAGAAATTTGA 2041

667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTGATACATTAGACCAAGCAAGAGTGTAAACTCCCAAGCTCTA 2091

684 ysSerGlyLysAsnPheSerTrpLeu 692
|||||
2092 AATCAGGTAAAAATTTCTCTTGGCTT 2117

seq_name: /net/abss06/STDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:AAx90695

seq_documentation_block:

ID AAx90695 standard; cDNA; 3058 BP.
XX AC AAx90695;
XX DT 01-OCT-1999 (first entry)
XX DE Human Cytoplasmic phosphatase Lyp 1, cDNA.
XX KW Lymphoid Protein Tyrosine Phosphatase gene; Lyp gene; lymphoid cell;
KW intracellular tyrosine phosphatase; PTPase; lymphocyte; thymocyte;
KW T cell; B cell; protein tyrosine kinase; PKs; immunosuppressant;
KW lymphoma; transgenic animal; T cell antigen receptor signalling;
KW cytokine receptor signalling; protein tyrosine phosphatase; PTPase; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 42..2468
FT FT /*tag= a
FT FT /product= "Cytoplasmic phosphatase, Lyp 1"
FT FT /note= "Expressed in thymocytes, mature T and B cells"
XX PN W09936548-A1.
XX XX
XX PD 22-JUL-1999.
XX XX
XX PF 18-JAN-1999; 99WO-CA00038.
XX XX
XX PR 16-JAN-1998; 98CA-2220853.
XX XX
XX PA (HSCR-) HSC RES & DEV LP.
XX XX
XX PI Roifman CM;
XX XX
XX DR WPI; 1999-444404/37.
XX DR P-PSDB; AAY28652.
XX XX

PT New nucleic acid encoding intracellular tyrosine phosphatase and
PT related proteins, used to modulate signalling through T cells,
PT particularly as immunosuppressant
XX
PS Claim 5a; Page 52; 105pp; English.

XX The present cDNA sequence is that of the Lymphoid Protein Tyrosine
CC Phosphatase gene, Lyp 1 that encodes a cytoplasmic tyrosine phosphatase
CC protein with a single catalytic domain. The Lyp gene has been localised
CC to human chromosome 1p13. It is expressed in the lymphoid cells,
CC particularly in thymocytes, mature B and T cells. It is involved in
CC lymphocyte growth and development. Lyp gene products are important for
CC regulation of T cell antigen and cytokine receptor signalling and early
CC and late stages of T cell differentiation. Lyp 1 has immunosuppressive
CC activity. This sequence is used to treat lymphoma, to reduce or prevent
CC T cell activation or proliferation and to control thymocyte
CC differentiation. Fragments of this gene are used, as probes or primers,
CC to identify allelic variations and for diagnosis of diseases associated
CC with reduced Lyp activity. Transgenic animals in which a Lyp gene has
CC been inactivated or replaced by a heterologous Lyp gene are used to
CC study and screen for possible therapeutic agents.

SQ Sequence 3058 BP; 1015 A; 592 C; 553 G; 898 T; 0 other;

alignment_scores:
Quality: 3573.00 Length: 685
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-600-358A-4 x AAx90695 ..

Align seg 1/1 to: AAx90695 from: 1 to: 3058

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLys 17
|||||
42 ATGGACCAAGAGAAATTCCTCAGAAAGTTCTGGATGAGGCCCAAGCAA 91

17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92 GAAATATTACTAAAGAGGAGTTTGGCAATGAATTTCTGAAGCTGAAAGGC 141

34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 ATCTACCAAGTACAAAGCAGACAAACCTATCTTACAACTGTGGCTGAG 191

51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 AATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATGA 241

67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242 TTATAGCCGGGTAGAACCTATCCCTGATAACCTCTGATGAGGATTCAGCT 291

84 YrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACCTTCATTAGGGAGTTTATGGACCCCAAGCTTATATT 341

101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIle 117
|||||
342 GCCACCCAGGGTCTCTTTATCTACAAACCTCTCTGGACTTCTGGAGGATGAT 391

117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTGTCTTATCATTTGTTATGCGATGCGATGGAGTATGAA 441

134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGATGCGAG 491

151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||

492 CTGGAATTTGGCCCTTCTCTGTATCTCTGAAAGCTCAAAAAAGGAATC 541
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgf 184
|||||
542 TGATTTATTAATCAGGACTCTAAAGTTAAGTTCAATAGTAGAACTCGAA 591
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTACCAAGTTTCATTAACAAGATTGGCCAGACCATGATGACCTTCA 641
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG1 217
|||||
642 TCTATAGACCTTATCTTGGAGCTCATCTGGGATGTCAGTTTGTACCAAGA 691
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgf 234
|||||
692 GGATGACAGTGTTCCTCATATGCTTCACTGCAGTGTGGCTGGGAAGGA 741
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTGTGCTATTGTTGATTATACATGCATGGATGTTGCTAAAGAT 791
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGATAATTCCTGAGAACTTCAGTGTGTTTCAGTTTGCATCGGGAATGCG 841
267 qThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCTTCATTAGTTCAACGCGAGGAACAATATGAAGTGGTCT 891
284 YrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACATGCTGTATTAGAACTATTAAAGAGACAGATGGATGTTATCAGAGAT 941
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTCGGAACAGAGAGTCAAGCAAGACATTTGTTCTCTCGAGAAAA 991
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAGCAGACTCTTATCTCTCTCTTAAATTTACCAAAAGTACCA 1041
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAAGCAGCAAAATGATGAACCAACAAAGGACAAAAATGGAATCAA 1091
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1092 GAATCTTCTCTCTTGTGACTTTAGGACTTCTGAAATAAGTGCAAAAGAGA 1141
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTGGCCCTGCTTAATCAACGACTCTCTTTGACTTCTCGGAGC 1191
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAAATTACAGTTTGGCAAAATGCTGCACCAACCATGAATGGCAGACA 1241
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAAATAGTTGGGAGCCTCTCTCAGAAGCATCAAAAGTTTGA 1291
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTTGTGGAGGATGTTCTAATCTTAAACCTGTAAATG 1341
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAAATTCAAAGTGCCAAATACACGGACCAATCA 1391
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCCTTTTGAATTTGATACAGAGAGAAACCAAGAGGTGGACAGCAA 1441

467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTTCTATTGGAATCTCAACCATGATCTTGTGTTGTAG 1491
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGAGGCTCAAAAAAGTAATGCATGTTTCTTTCAGCAGAACTGAATTA 1541
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACTGCCATATGACTTAACACCAAAATACGTAATGCCCTTAATGFAAA 1591
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
1592 GCACCATGACTCTAGTCTCTTGGTGTATATCTTACATACCTTTAGTGG 1641
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGCTCCAAGTGGTACCAGTTCTTAAGATG 1691
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTACCTGAGAAAGCAAGATGGAACCTGTTTCTCTCTCTCT 1741
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCCAACATCTCTACATCCCTCTCTCTTATTACAAATCACAATGTT 1791
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCAGTGAATTTCTCAACCAATATTTCTCCTACATTTGAACCCAGAG 1841
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCACTGCTCCAAGGATAGATGATGAATCCCTCCCTCC 1891
617 oLeuProValArgThrProGluSerPheIleValValGluAlaGlyG 634
1892 ACITCTGTAGGACACCTCAATCATTTATTTGTTGGTGGAGAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAATGTTCCCAAAATCTTATCTCAGCTGTGAAGGTAAA 1991
651 IleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAACATCAGTGGAAATGGGTGGAAACATCTGAACCAAGAAATTTGA 2041
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTATCTAGTACCAACAGAGTGTAAACTCCGAAGTCCCTA 2091
684 ysSer 685
|||||
2092 AATCA 2096

seq_name: /net/abss06/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:AAV81742

seq_documentation_block:

ID AAV81742 standard; cdNA; 3580 BP.

XX AAV81742;

AC

XX 10-MAR-1999 (first entry)

DT Human PTP04 encoding cdNA.

XX

XX PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;

XX type I receptor serine/threonine kinase; cancer; leukemia; lymphoma;

XX neurodegenerative disease; neuronal survival; Alzheimer's disease;

XX Parkinson's disease; Huntington's disease; ss.

XX

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Key Location/Qualifiers
 FH 53..2476
 FT CDS /*tag= a
 XX

PN W09849317-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1100 GAATCTCTCTCTTGCATCTTAGCACTTCTGAATTAAGTGCAAAAGAGA 1149
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1150 GCTAGTTTTCACCCCTGCTAAATCAAGCACTTCTTTTGACTTTCTGGAGC 1199
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1200 TAAATACAGTTTGTGACAAATATGCTGACACCAACCATGAATGGCAGACA 1249
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1250 AAGGCATTTCCAAATAGTTGGGAGCCTCTTCAGACGATCAAAAGTTTGA 1299
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1300 TTTGGGCTCTCTTTTGTGAGGATGTTCTTAATTTCTAAACCTGTAAATG 1349
434 laalaglyargTyrPheAsnSerLysValProIleThrargThrLysSer 450
|||||
1350 CACGAGAAGATATTTTAATTCAAAGGTGCCAATAACACGACCAATCA 1399
451 ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
|||||
1400 ACTCCTTTGAAATGTATACAGACAGAGAAACCAAGAGGTGGACAGCAA 1449
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1450 GGAATACTTTCTTATTTGGAATCTCAACACACATGATTTCTTTGTAG 1499
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1500 AGATGAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAGCAACTGAATTA 1549
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1550 TCAGTCCATATGACTTAACACCAAAATACGTAAATGCCTCTAATGTAAA 1599
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1600 GCACCATGACTCTAGTCTCTTGGTGATATTTCTTACATACCTTTAGTGG 1649
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1650 AAATCCTTATTTTTCATCATGCGCTCCCAAGTGGTACCAGTTCTAAGATG 1699
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1700 TCTCTTGATTTTACCTGAGAGCAAGATGGAACTGTTTTCCTTCTTCTCT 1749
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1750 GTTGCCAAACATCTCTACATCCCTCTCTCTTCTTATTAACAATTCATGATT 1799
584 erLeuSerLeuAsnSerProThrAsnIleSerSerSerLeuLeuAsnGlnGlu 600
|||||
1800 CTTTATCAGTGAATTTCCAAACCAATATTTCCCTCACTATTGAACCCAGGAG 1849
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1850 TCAGCTGTACTAGCACTGCTCCAGGATAGATGAAATCCCCCTCC 1899
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1900 ACTTCTGTACGGACACCTCAATCATTTATTGTGGTTGAGGAAGCTGGAG 1949
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1950 AATTCTCACCAAAATGTTCCCAAAATCTTATCTCAGCTGTGAAGGTAATA 1999

651 ileGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
2000 ATTGGAACATCACTGGAATGGGTGGAACATCTGAAACCAAGAAATTTGA 2049
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2050 TGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAAACTCCGAAGTCTTA 2099
684 ysSer 685
|||||
2100 AATCA 2104
seq_name: /net/abss06/sids1/gcgdata/hold-geneseq/geneseq-n-emb1/NA2001A.DAT:AAH79342
seq_documentation_block:
ID AAH79342 standard; DNA; 2452 BP.
XX
AC AAH79342;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase coding sequence.
XX
KW Human; tyrosine phosphatase; hPTP; ds.
XX
OS Homo sapiens.
XX
PN CN1302899-A.
XX
PD 11-JUL-2001.
XX
PF 29-OCT-1999; 99CN-0119935.
XX
PR 29-OCT-1999; 99CN-0119935.
XX
PA (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
PI Ren S, Wu T, Qian B;
XX
DR WPI; 2001-550591/62.
DR P-PSDB; AAG78623.
XX
Human protein tyrosine phosphatase and its coding sequence -
PS Claim 1; Page 19-21(Disclosure); 29pp; Chinese.
XX
CC The present invention provides the protein and coding sequences of human
CC tyrosine phosphatase hPTP. The protein is expressed in human normal
CC suprarenal tissue. The present sequence is the coding sequence of the
CC invention.
XX
SQ Sequence 2452 BP; 791 A; 516 C; 468 G; 677 T; 0 other;

alignment_scores:
Quality: 2977.50 Length: 687
Ratio: 4.660 Gaps: 11
Percent Similarity: 93.013 Percent Identity: 89.229
alignment_block:
US-09-600-358A-4 x AAH79342 ..
Align seg 1/1 to: AAH79342 from: 1 to: 2452
1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
10 ATGGACCAAGAGAAATTTCTGCAGAAAGTTCTCGATGAGGCCCAAGCAA 59
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
|||||
60 GAAATTAATAAGAGAGAGTTTCCCAATGAATTTCTGAAGCTGAAAGGC 109
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrValAlaGlu 50

|||||
110 AATCTACCAAGTACAAGGCAGACAAAACCTATCTACAACTGTGGCTGAG 159
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
:::
160 AAGCCCAAGAAATATCAGAAAAACAGATATAAGGATATTTGGCCCTATGA 209
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
210 TTATAGCGGGTAGAACTATCCCTGATAAACCTCTGATGAGGATCCAGCT 259
84 yrlleAsnAlaAspPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
260 ACATCAATGCCAACTTCATTAAAGGAGGTTTATGACCCGAGGCTTATATT 309
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetI1 117
310 GCCACCCAGGGTCTTTATCTACAAACCTCTCTGGACTTCTGGAGGATGAT 359
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
360 TTGGGAATATAGTGCCTTATCATTTATGTCATGTCATGATGAGTATGAAA 409
134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
410 TGGGAAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGGAGATGCGAG 459
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
460 CTGGAAATTTGGCCCTTTCTCTGTATCTGTGAAGCTGAAAGAAAGCAATC 509
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
510 TGATATATATATAGGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 559
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
560 CTATCTACCACTTTCATTACAAAGATTGGCCAGACCATGATGATACCTTCA 609
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG1 217
610 TCTATAGACCCCTATCTTGAGCTCATCTGGGATGTCCGTTGTACCCCAAGA 659
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
660 GGATGACAGTGTCCCATATGCAATTCATCTACTGCAGTCTGGCTGTGGAAGA 709
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
710 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGCTTAAAGAT 756
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
757 GGGATAATCTCTGAGAACTTCAGTGTGTTTTCAGTTTGATCCGGGAAATGCG 806
267 qThrGlnArgProSerLeuValGlnThrGlnGlnTyrGlnLeuValT 284
807 GACACAGAGGCTTCATTAGTTCAACGCAGGACAATATGAACCTGGTCT 856
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
857 ACAATGCTGATTAGAACTATTAAAGACAGACAGATGGATGTTATCAGAGAT 906
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
907 AAACATTCTGGAACAGAGAGCTCAAGCAAGCATTTGATTCCTGAGAAAAA 956
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
957 TCACACTCTTCAACGACACTCTTATCTCTTAATTTACCAAAAAAGTACCA 1006
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||

1007 CAAAGCAGCAAAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1056
351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1057 GAATCTCTCTCTTGTGACTTTAGGACTCTGAAATAAGTGCAAAAGGAAGA 1106
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
1107 CGTAAGTTTGGCCCG...GAAACACAGGCTTTTTCCTTCCTTCGGAAG 1153
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
1154 TAAATTCAGTTTGCAAAATGTTG...CCCCACCATGAATGCCAGCA 1200
401 LysAlaPheProIleValGlyGluPro LeuGlnLysHisGlnSerLeuA 417
|||||
1201 AGGCTTTCNAATAGTGGGA...AGCTCTTCAGAGCATCAAAAT...T 1244
417 spLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsn 433
1245 GAATTTGGGCTCTCTTGTGTTAGGAATGTCTA...ATTCTAACCTTGAT 1291
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
1292 GCAGCAGAAAGATATTTAATCCAAGGTGCCAAT...ACCCGACCAATC 1338
450 rThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerL 467
1339 AACCTT.TTTGAATTGATCCAGCAA...GAGAACCAAGGAGGTGACAGCA 1384
467 ysGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheVal 483
1385 AGAAACTTCTTATTT...GAATCTCAACCATGATCTCTGTTTGTGA 1431
484 GluMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTy 500
1432 GAGATGAGGCTCAAAAAGTAATGATGATTTCTTCAGCAGAACTGAATTA 1481
500 rSerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVall 517
1482 TTCCTGCCATATGACTCTAAACCAACCAATACGTAATGCCCTCTAATGTAA 1531
517 ysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVal 533
1532 AGCACCATGACTCTAGTGTCTTGTGTATATTCTTACATACCTTTAGTG 1581
534 GluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMe 550
1582 GAAATCTCTTATTTTCATCATGTGGCTCCAAAGTGTACCAAGTTCTAAGAT 1631
550 tSerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerL 567
1632 GTCTCTGATTTACTGAGAACCAAGATGGAACGTGTTTCTCTCTCTCTC 1681
567 euLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSer 583
1682 TGTGTCACCAATCTCTACATCCCTCTCTCTTATTACAAATTCACATGAT 1731
584 SerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnG1 600
1732 TCTTTATCACTGAATTTCTCCAACCAATATTTCTCAGGTATTGAACAGA 1781
600 uSerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProp 617
1782 GTCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATGAAT.CCCCTC 1830
617 roLeuProValArg.ThrProGluSerPheIleValValGluGluAlaG1 633
1831 CACTTCTGTACGGGACACCTGATCATATTATTGTTGAGGAGGCTGG 1880
633 yGluPheSerProAsnValProLysSerLeuSerSerAlaValLysVall 650
1881 AGAATTCACCAAAATGTTCCCAAAATCCTTATCTCAGCTGTGAAGTAA 1930
|||||


```
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
|||||
899 ATGAACTTGTATAGGGCTATGCTCAACTGTTGAAACAGCTACAC 948
297 ValIle..... 298
949 TGTATGAATTCATGAGCAGACAGAAGATCCGTGATGGTAATGAAATTACC 998
299 .....ArgAspLysHisSer. 303
999 ACTGGAACATATGCTAGTTCCATTGATAGCGAGACGACCTCTCCGCCA 1048
304 .....GlyThr 305
1049 AAGCGCGCGGACTCGAAGTTGCCCTTGTAGAAGGGGATGCCAAGGAAGA 1098
306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
1099 AATACTACAGCCACCAGAACCTCACCCGGGTGCCACCCCTCTGAGGCCAT 1148
322 aAspSerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysM 339
1149 CACCTCCCTTCACG.....CTTCCCAACGTTACCACTGTGTGCAGGACAG 1192
339 et.....MetAsn 341
1193 TCAGAGGTACACCCAAAGCCAGTCTGCACATGCGCTCACCAGAGCAAC 1242
342 GlnGlnArgThrLysMetGluIleLysGluSerSerPheAspPheAr 358
1243 ACCCAGCGCCACTCAACAGAA.....GCTATGATAAATCAGCG 1280
358 qThrSerGluIleSerAlaLysGluLeuValLeuHisProAlaLysS 375
1281 GACC.....AATGGGCAAAAA 1297
375 exSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
1298 GTGAATCAGCTATTGAGCACATA.....GATAAGAG 1329
392 AlaAspThrThrMetLysTrpGlnThrLysAlaPheProIle..... 405
1330 TTAGAGCGCAATTAAAGTTTTCAGATTAAAGAAAGTCCCTCTCCAAGAAG 1379
405 ..... 405
1380 GCCCAAAAGTTTGTAGGGAACACACTCTTGAATAGGGGACATGCGATTA 1429
406 .....ValGlyGluProLeuGlnLysHisGln 414
1430 AAATTAATATCTGCTTCATCTTCTGTAGTTGACAGAACCTCTAAACACAG 1479
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
1480 GAGTTAAGTCGAGGTGGCCCTAAAGTTGATGTATCTCAGAAATCTTG 1529
431 oValAsnAlaAlaGlyArgTyr..... 438
1530 CCGGACTGTAGTGGCGCTCATTCACAGAGCTGCTGAGTCGTGAGG 1579
439 .....PheAsnSerLysValProIleThrArgThrLysSerThrPro 452
1580 AGTCCCGAGCACTACACACACT...CCACGGCCAGACTGCTTGCT 1626
453 PheGluLeuIleGlnArgGluThrLysGluValAspSerLysGluAs 469
1627 CTCGATAAG...AAAGGACAGTAACGTGGTCACTTCATGACCTGAAAA 1673
469 nPheSerTyrLeuGluSerGlnProHisAspSerCysPheValGluMetG 486
1674 TGCCACACTGTATCCCGACTCACCT...GACGGCAAAATCCCCAGATAATC 1720
```

```
486 InAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeu 502
|||||
1721 ATTCTCAGACTCTGAAACCGGTGAGTTCCACACC.....AACTCCACC 1764
503 ProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisH 519
1765 GCAGAGAGGAGGAGCCCGACGATCTT.....ACAGAGCACCA 1799
519 sAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluAsn 536
1800 CAACAGCTCC.....CCTCTGTTGAAAGCTC 1825
536 to.....TyrPheSerSerTyrProProSerGly 545
1826 CCTCAGCTTTACCAACCTCTTCACTCTGAGGACTGCGACTCAGACGA 1875
546 ThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrVa 562
1876 GGGAGCTCT.....GATGGTGTCTGT 1895
562 lPheProSerSerLeuProThrSerSerThrSerLeuPheSerTyrT 579
1896 GACCAGGACAAACTAGCATTTCAACAGCAAGTCCACAGTGTCTCTCTG 1945
579 yrAsnSerHisSerSerLeuAsnSerProThrAsnIleSerSer 595
1946 CCAGTAGTCTGAGAGT...GCTTCCATAGGAGAGTATGCCGATGTC 1992
596 LeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArg..... 609
1993 ATTGCCAGACAGGAA.....GTAGCAGGACGCGCATTCAGGTGC 2033
610 .....IleAspAspGluIleProProProLeuProVala 621
2034 TCAGAAAGATGCTGATGTTAGTAGGAGTGCCTCTCTCTTACCTGAAC 2083
621 rgThrProGluSerPheIleValValGlu 630
2084 GAACCTCTGAGCTTTTCTGTTATAGCAGAT 2112
```

seq_name: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.AAF21801

seq_documentation_block:

ID AAF21801 standard; DNA: 1286 BP.

XX AAF21801;

XX 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 188.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
XX nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; anticancer; vulnery; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
XX Addison's disease; allergy; autoimmune haemolytic anaemia;
XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis; ds.
XX cardiovascular disorder; wound healing; neurological colitis; ds.

XX Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX
DR P-PSDB; AAB58898.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 1; Page 622; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 1286 BP; 410 A; 273 C; 209 G; 389 T; 5 other;

alignment_scores:
Quality: 1058.00 Length: 208
Ratio: 5.111 Gaps: 0
Percent Similarity: 99.519 Percent Identity: 99.038

alignment_block:

US-09-600-358A-4 x AAF21801 ..

Align seg 1/1 to: AAF21801 from: 1 to: 1286

478 HisAspSerCysPheValGluMetGlnAlaGlnLysValMetHisValSe 494
|||||
2 CATGATCTCTGTTTGTAGAGATCGAGCTCAAAAGTAATGCAITTC 51
494 rSerAlaGluLeuAsnTyrSerLeuProTyrAspSerLysHisGlnIleA 511
|||||
52 TTCAGCAGAACTGAATTATTTCACCTGCCATATGACTCTAAACACCAATAC 101
511 rGAsnAlaSerAsnValLysHisAspSerSerAlaLeuGlyValTyr 527
|||||
102 GTAATGCTCTTAATGTAAGCACCACCATGCTAGTGTCTTGTGTATAT 151
528 SerTyrIleProLeuValGluAsnProTyrPheSerSerTrpProProSe 544
|||||
152 TCTTACATACCTTTAGTGAAATCCCTATTTTCATCATGGCCTCCAAG 201
544 rGlyThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyT 561
|||||
202 TGGTACCAGTTCTTAAGATGTCTCTTGATTTTACCTGAGAAGCAAGATGAA 251
561 hrValPheProSerSerLeuLeuProThrSerSerThrSerLeuPheSer 577
|||||
252 CTGTTTTTCTCTCTCTGTGTGCAACATCTCTATACATCCCTCTCTCT 301
578 TyrTyrAsnSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSe 594
|||||
302 TATTACAATTCACATGATCTTTATCACTGAATTCCTCAACCAATATTC 351
594 rSerLeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArgIleA 611

|||||
352 CTCACCTATTGACCCAGGAGTCAGCTGTACTAGCAACTGCTCCAGGATAG 401
611 spAspGluIleProProLeuProValArgThrProGluSerPheIle 627
|||||
402 ATATGAAATCCCCCTCCACTTCCCTGTAGCGACACCTGAATCATTTATT 451
628 ValValGluGluAlaGlyGluPheSerProAsnValProLysSerLeuSe 644
|||||
452 GTGGTTGAGGAAGCTGGAGAAATCTCACCAGAAATGTTCCCAATCCTTATC 501
644 rSerAlaValLysValLysIleGlyThrSerLeuGluTrpGlyGlyThrS 661
|||||
502 CTCAGCTGTCAAGTAAATTTGGAACATCCTGGAATGGGTGGAACAT 551
661 erGluProLysLysPheAspSerValIleLeuArgProSerLysSer 677
|||||
552 CTGAACCAAGAAATTTGATGACTCTGTGATCTTAGACCAAGCAAGAGT 601
678 ValLysLeuArgSerProLysSer 685
|||||
602 GTAAACTCCGAAGTCTTAATCA 625

seq_name: /het/abss06/SIDS1/gcgdata/hold-geneseg/geneseqn-emb1/NA1998.DAT:AAV17097

seq_documentation_block:

ID AAV17097 standard; cDNA; 2226 BP.

AC AAV17097;

DT 20-JUL-1998 (first entry)

DE Rat protein tyrosine phosphatase PRP20 cDNA.

KW Protein tyrosine phosphatase; PTP20; rat; signal transduction;

KW cell differentiation; cancer; neural injury; therapy; ss.

OS Rattus sp.

FH Key Location/Qualifiers

FT CDS 28..1389

FT polyA_signal /tag= a

FT WO9748723-A2. 2150..2156

XX /tag= b

XX 24-DEC-1997.

XX 17-JUN-1997; 97WO-IB00946.

XX 19-DEC-1996; 96US-0034286.

XX 17-JUN-1996; 96US-0019629.

XX 09-AUG-1996; 96US-0023485.

XX 13-NOV-1996; 96US-0030860.

XX 15-NOV-1996; 96US-0030964.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Aoki N, Chen Z, Kharitonov AI, Kim YW, Nayler O;

XX Ullrich A, Wang HY;

XX WPI: 1998-120302/11.

XX P-PSDB; AAW49906.

XX New phosphatase and kinase enzyme(s) - useful in the diagnosis and

XX treatment of signal transduction disorders

XX Claim 3; Fig 1a-b; 138pp; English.

XX This cDNA clone codes for a novel rat protein tyrosine phosphatase,

XX designated PTP20 (see AAW49906), that regulates growth factor

XX stimulation of cellular differentiation. The clone was isolated

CC from a rat phaeochromocytoma PC12 cDNA library by PCR amplification
 CC using primers based on consensus sequences (see AAW49915-16) of known
 CC ptps. The invention relates to novel proteins (see AAW49906-14)
 CC involved in cellular signal transduction and to the nucleic acids
 CC (see AAV17097-99) coding for them, and provides vectors, host cells,
 CC purified recombinant proteins, methods for identifying compounds
 CC that bind the novel proteins or abrogate their interactions with
 CC natural binding partners, as well as methods for the diagnosis and
 CC treatment of diseases associated with the novel proteins. For
 CC ptp20, activators may act as anti-cancer therapeutics that
 CC stimulate cell differentiation rather than proliferation, while
 CC inhibitors may be useful for treating neural injuries by delaying
 CC the differentiation of transplanted neuronal stem cells until they
 CC are firmly grafted.

XX
 SQ Sequence 2226 BP; 574 A; 591 C; 604 G; 457 T; 0 other;

alignment_scores:
 Quality: 748.00 Length: 357
 Ratio: 3.041 Gaps: 7
 Percent Similarity: 68.908 Percent Identity: 43.137

alignment_block:

US-09-600-358A-4 x AAV17097 ..

Align seg 1/1 to: AAV17097 from: 1 to: 2226

```

3 GlnArgGluLeuGlnLysPheLeuAspGluAlaGlnSerLysLysI 19
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
37 CAATCGGACCTAGTGAGGAGCTCTCTGGAGCAGCAGGAGCGCGGACCA 86
19 eThrLysGluGlu...PheAlaAsnGluPheLeuLysLysLysArgGlns 35
||| :|||:|||||:|||||:|||||:|||||:|||||:
87 CGGGAAGGGGGCAATCCCTCGCCGGTGGTTCAGCAGCAATTAAGGCCCGCT 136
35 eThrLysLysLysAlaAspLysThrTyPrThrValAlaGluAsn 51
|||: :|||:|||||:|||||:|||||:|||||:|||||:
137 CAGTGGCTTGAAGACTGAAGGTGTGCTCCACTTAAGCGCGCAGTCAG 186
52 AlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAspTy 68
||| :|||:|||||:|||||:|||||:|||||:|||||:
187 CAGGGAACCTCAAGAAGAACCGCTACAAAGACGTGTACCGCTATGATGA 236
68 rSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSeryrI 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:
237 GACGAGGTGTCATCCCTTCCCTGCTCCAGGAGGAGGACACCGGAGATTACA 286
85 LeAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIleAla 101
|||:|||||:|||||:|||||:|||||:|||||:|||||:
287 TTAATGCCAACTTCATCCGGGCACAGATGGNAGCCAGCCCTACATTCGG 336
102 ThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTr 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 AGCGAAGGACCCCTGCCCTCACACTCTGTGGACTTCTGGCGCTGTGTTG 386
118 pGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluMetG 135
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 GGAGTTTGAATCAAGGTGATCTTGTGCTGCTGTCAGGAGCAGAAAAATG 436
135 LyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu 151
|||:|||||:|||||:|||||:|||||:|||||:|||||:
437 GACGGAGGAAGTGTGAAGCTACTGGGCCCGCCAGGACGGGAG...CCTCTA 483
152 GluPheGlyProPheSerValSerCysGluAlaGluLysArg...LysSe 167
|||:|||||:|||||:|||||:|||||:|||||:|||||:
484 CAGGCGCGGCGCTTCTGTCATCACCTGACAAAGGAGAGCAGCACTGACTTC 533
167 rAspTyrIleLeuArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
534 GGACATCACTCTCAGGACCCCTCCAGGTTTACATTTCCAGAAGGAATCCCGTC 583
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200

```

```

584 CTGTGCACCAGCTACAGTACATGTCTTGGCGGACCACCGGGTTCCACAGC 633
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnG 217
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
634 AGTTCCGATCATCTTCACCATGGTGGAGGAGCCCGTGCCTCCAAAGG 683
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
684 ACTTGGACCTGGACCCCTCTGTGCTCCACTGAGTGTGGCTGGCTGGACGAA 733
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
734 CAGGTGTCTGTGTGCT...GTTGATTACGTGAGGCAGTGTCTTCTGACT 780
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781 CAGCAATCCCAACCAATTCAGCCTCTTGAAGTGGTCTGGAGATGCG 830
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
831 GAACAGGACCTGCGCGGTGCAGACAGAGGAGGAGGAGTACAGGTTCCTGT 880
284 yrAsnAlaValLeuLeuPheLysArgGlnMetAsp..... 296
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
881 ACCACACAGTGGCTCAGCTATTCTCCCGCACTCTCCAGAACACACAGTCCC 930
297 .....ValIleArgAspLysHi 302
931 CTCTACCAAGAACCTCAAGGAGAACCGCGCTCCAATCTCAAGGATTCCTC 980
302 sSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHisT 319
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
981 GTCCCTCAGGACCTCCTCAGCC.....CTGCCTGCCACATCCCGCC 1021
319 hrLeuGlnAlaAspSerTyrSer.....ProAsnLeu 329
||| :|||:|||||:|||||:|||||:|||||:|||||:
1022 CACTGGGTGGGTCTCTCAGGAGCATCTCGTGTGCTGGCCACCGACCTT 1071
330 ProLysSerThrThrLysAla 336
||| :|||:|||||:|||||:|||||:|||||:|||||:
1072 CCCATGGCTGACACTTACGCT 1092

```

seq_name: /net/abss06/SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:AAV03112

seq_documentation_block:

ID AAV03112 standard; cDNA; 1529 BP.

XX
 AC AAV03112;

DT 09-APR-1998 (first entry)

XX cDNA encoding a novel non-receptor tyrosine phosphatase.

DE Non-receptor protein tyrosine phosphatase; hematopoietic stem cell;
 KW PTP HSC; progenitor cell; tyrosine phosphatase domain; PTP HSC agonist;
 KW tyrosine dephosphorylation; tyrosine phosphatase; PTP HSC antagonist;
 KW stem cell differentiation; ss.

OS Mus sp.

PH Key Location/Qualifiers

FT CDS 19..1380

FT /*tag= a

XX WO9735019-A1.

XX 25-SEP-1997.

XX 17-MAR-1997; 97WO-US05278.

XX 22-MAR-1996; 96US-0620526.

XX

(GETH) GENENTECH INC.

Cheng J, Lasky LA;

WPI; 1997-480224/44.

P-PSDB; AAW37254.

Hematopoietic stem cell non-receptor protein tyrosine phosphatase -
useful for expansion of undifferentiated stem cells in cell culture

Disclosure; Fig 1; 66pp; English.

The present sequence encodes a novel murine non-receptor protein tyrosine phosphatase of hematopoietic stem cells (PTP HSC). This protein is predominantly expressed in early hematopoietic stem cells (HSCs) or progenitor cells, and lacks expression in adult tissues. The protein has a N-terminal tyrosine phosphatase domain, followed by a region rich in serine, threonine and proline and a C-terminal region of about 15-25 amino acids which is rich in basic amino acid residues. The protein is capable of tyrosine dephosphorylation in hematopoietic progenitor cells, and functional derivatives of such native tyrosine phosphatases. The phosphatase domain of the PTP HSC or a PTP HSC-expressing HSC or progenitor cell can be used in an assay for the identification of PTP HSC antagonists or agonists. The antagonist can be used to induce the differentiation of stem cells, such as undifferentiated malignant hematopoietic cells, e.g. leukemia cells, which may facilitate their treatment. The PTP HSC or an agonist antibody against the PTP HSC can be used for the expansion of undifferentiated stem cells in cell culture (this allows expansion of HSC prior to autologous or heterologous bone marrow transplantation), while the agonist antibody along with a hematopoietic growth factor can be used for the expansion of undifferentiated stem cells in vivo.

Sequence 1529 BP; 358 A; 437 C; 445 G; 289 T; 0 other;

alignment_scores:

Quality: 736.50 Length: 327
Ratio: 3.069 Gaps: 6
Percent Similarity: 73.394 Percent Identity: 44.037

alignment_block:

US-09-600-358A-4 x AAV03112 ..

Align seg 1/1 to: AAV03112 from: 1 to: 1529

```

5  GluLeuGlnLysPheLeuAspGluAlaGlnSerLysLysLysLysLysL 21
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34  GACTTGTGAGGAGCTTCTTGGAGCAGCTGGAGGCGCGGACTACCGGGA 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
21  sGluGlu...PheAlaAsnGluPheLeuLysLeuLysArgGlnSerThrL 37
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84  GGGGGCAATCTTCGTCGTGAGTTCACGACATTAAGGCCGCTCAGTGG 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37  ySTyRysAlaAspLysThrProThrThrValAlaGluAsnAlaLys 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
134  CCTGGAAGTCTGAAGGTGTGTGCCACTAAAGCCGCGAGCTGGCTTGGG 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54  AsnLleLysLysAsnArgTyRysAspLleLeuProTyRasPyrSerAr 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184  AACACGAACAAAGAACCCCTCAAGATGTGTGATGATGATGAGCAAG 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
70  qValGluLeuSerLeuLysThrSerAspGluAspSerSerTyRileAsnA 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234  AGTCATCTTCTCCCTGCTCAAGAGGAGGACATGGAATTCATCAATG 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
87  laAsnPhelleLysGlyValTyRglyProLysAlaTyRilleAlaThrGln 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284  CCAACTTCATCCGGGGCATAGATGGAAGCCAGGCCCTACATTCGACGCAA 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
104  GlyProLeuSerThrThrLeuLeuAspPheThrArgMetIleTrpGluTy 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334  GGACCCCTCTCACACACTGTTGGAGTCTTGGCGCTGTTGGGAGTT 383

```

```

120  rSerValLeuIleValMetAlaCysMetGluTyRgLuMetGlyLysL 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384  TGGGGTCAAGGTAAATCTGATGCCCTGTCAAGAGACAGAAAATGGACGGA 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137  ySLyScysGluArgTyRTrpAlaGluProGlyGluMetGlnLeuGluPhe 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434  GGAAGTGTGAACGCTATTGGCCGCGGAGCAGGAG...CCTCTAAAGGCT 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
154  GlyProPheSerValSerCysGluAlaGluLysArg...LysSerAspTy 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
481  GGGCCCTTCTGTCATCCCTTGACAAAGGAGACACATGTAATCGACACAT 530
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
169  rLleIleArgThrLeuLysValLysPheAsnSerGluThrArgThrIle 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
531  CACTCTCAGACCCCTCCAGGTTACATTCACAGAAATTCGGCTCTGTGC 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
186  yRGlnPheHisTyRysAsnTrpProAspHisAspValProSerSerIle 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
581  ACCAACTACAGTATATGTCCTGCCAGACCACGGGTTCACGACAGTTCT 630
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203  AspProIleLeuGluLeuIleTrpAspValArgCysTyRglnGluAspAs 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
631  GATCATTCTTCCATGTTGGAGGAGGCCGCTGCCCTCCAAGGGCTGG 680
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
219  pSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGlyV 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
681  ACCTGGACCCCTCTGCTCCTCAGTGCCTGGCTGGGAGCAACAGGTG 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236  alLeCysAlaIleValAspTyRThrTrpMetLeuLeuLysAspGlyIle 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
731  TCCTGTGCGCT...GTTGACTATGTGAGCGAGTGTCTGTGACCCAGACA 777
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
253  IleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThrGl 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
778  ATCCCTCCCAACTCACTCTTCTCCAGTGGTCTGGAGATGCGGAAACA 827
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
269  nArgProSerLeuValGlnThrGlnGlnTyRglnLeuValTyRAsnA 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
828  GCGGCTGCAGCAGTGCAGACAGAGGAGCAGTACAGGTTCTGTACCACA 877
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
286  laValLeuGluLeuPheLysArgGlnMetAspValIleArgAspLysHis 302
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
878  CAGTGGCTCAGTATTCTCCGCG...ACTCTCCAGGACACCCAGC 918
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
303  SerGlyThrGluSerGlnAlaLysHisCysIlePro.....GluLysAs 317
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
919  CCCCATTACCAAGACCTCAAGGAGAACTGCGCTCCAATCTGCAAGGAAGC 968
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317  nHisThrLeuGlnAlaAspSerTyRSerPro 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
969  TTTCTCCCTCAGGACCTCTCCTCAGCCCTGCCT 999
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

seq_name: /net/abss06/SIDSL/gcgdata/hold-geneseg/geneseqn-emb1/NA1998.DAT:AAV17099

seq_documentation_block:

ID AAV17099 standard; cDNA; 2810 BP.

XX AAV17099;

XX 20-JUL-1998 (first entry)

XX Human brain derived phosphatase 1 (BDP-1) cDNA.

XX Brain derived phosphatase 1; BDP-1; human; receptor;
 KW protein tyrosine phosphatase; signal transduction;
 KW therapy; diagnosis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH misc_feature 31..43

FT /*tag= a

```

FT CDS /note= "GC-rich track, part of Kozak sequence"
FT 44..1420
FT /*tag= b
FT /transl_except= (pos:69..71, aa:Arg)
FT /transl_except= (pos:162..164, aa:His)
FT /transl_except= (pos:1182..1184, aa:Ser)
FT intron 733
FT /*tag= c
FT /note= "site of 367 bp intron"
FT intron 799
FT /*tag= d
FT /note= "site of 80 bp intron"
FT intron 878
FT /*tag= e
FT /note= "site of 91 bp intron"
FT misc_feature 2532..2545
FT /*tag= f
FT /note= "t-rich sequence required for
FT polyA_signal 2740..2747
FT /*tag= g
FT WO9748723-A2.
FT 24-DEC-1997.
FT 17-JUN-1997; 97WO-IB00946.
FT 19-DEC-1996; 96US-0034286.
FT 17-JUN-1996; 96US-0019629.
FT 09-AUG-1996; 96US-0023485.
FT 13-NOV-1996; 96US-0030860.
FT 15-NOV-1996; 96US-0030964.
FT (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
FT Aoki N, Chen Z, Kharitonov AI, Kim YW, Nayler O;
FT Ullrich A, Wang H;
FT WPI: 1998-120302/11.
FT P-PSDB: AAW49908.
FT New phosphatase and kinase enzyme(s) - useful in the diagnosis and
FT treatment of signal transduction disorders
FT Claim 3; Fig 3a-d; 138pp; English.
FT This cDNA sequence codes for a novel human protein tyrosine
FT phosphatase (PTP), designated brain derived phosphatase 1 (BDP-1,
FT see AAW49908), that is expressed in most tissues and cell lines at
FT basal level, but expressed high in epithelium origin cell lines
FT and cancer cell lines. BDP-1 was originally identified in a human
FT brain cDNA library, although the full-length clone was isolated
FT from the haematopoietic MEG01 cDNA library. The invention relates
FT to novel proteins (see AAW49906-14) involved in cellular signal
FT transduction and to the nucleic acids (see AAV17097-99) coding for
FT them, and provides vectors, host cells, purified recombinant
FT proteins, methods for identifying compounds that activate or
FT inhibit the novel proteins, as well as methods for the diagnosis
FT and treatment of diseases associated with the novel proteins.
FT Sequence 2810 BP; 700 A; 808 C; 808 G; 494 T; 0 other;

alignment_scores:
  Quality: 708.00 Length: 291
  Ratio: 3.218 Gaps: 4
  Percent Similarity: 75.601 Percent Identity: 46.735

alignment_block:
  US-09-600-358A-4 x AAV17099 ..
  Align seg 1/1 to: AAV17099 from: 1 to: 2810

```

```

25 AlaAsnGluPheLeuLysLeuLysArgGlnSerThrLysTyrLysAlaAs 41
|||||
122 GCGGCGAGTTACAGGACATCCAGGCGCTCTCGCGCGCTGGAGGCTGA 171
|||||
41 pLysThrTyrProThrThrValAlaGluAsnAlaLysAsnLysLysA 58
|:::
172 CGGCGTGTCTCCACCGTGGCGGCGAGTCGCCAGAGACGTGAGGAAGA 221
|||||
58 snArgTyrLysAspIleLeuProTyrAspTyrSerArgValGluLeuSer 74
|||||
222 ACCGCTACAAAGACGTGCTTATGATCAGACGCGAGTAATCTCTCC 271
|||||
75 LeuIleThrSerAspGluAspSerSerTyrIleAsnAlaAsnPhellely 91
|||||
272 CTGCTCCAGGAGAGGACACAGCGACTACATTATGCAACTTCATCCG 321
|||||
91 sGlyValTyrGlyProLysAlaTyrIleAlaThrGlnGlyProLeuSer 108
:|||||
322 GGGCGTGGATGGAAGCCTGGCTTACATTGCCACGCAAGGACCTTGCCTC 371
|||||
108 hrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrSerValLeuIle 124
|||||
372 ACACCTGCTAGACTTCTGGAGACTGGGTCTGGGAGTTTGGGGTCAAGGTG 421
|||||
125 IleValMetAlaCysMetGluTyrGluMetGlyLysLysLysCysGluAr 141
|||||
422 ATCTGTATGGCTGTGAGAGATAGAGAAATGGCGGAAAAGGTGTAGCGG 471
|||||
141 gTyrTrpAlaGluProGlyGluMetGlnLeuGluPheGlyProPheSerV 158
|||||
472 GTACTGGGCCAGGAGGAGGAG...CCACTGCAGAGTGGGCTTTCTGCA 518
|||||
158 alSerCysGluAlaGluLys...ArgLysSerAspTyrIleIleArgThr 173
|||||
519 TCACCTCTGATAAAGAGAGAGTGGCTGAATGAGGACATCATGCTCAGGACC 568
|||||
174 LeuLysValLysPheAsnSerGluThrArgThrIleTyrGlnPheHisty 190
|||||
569 CTCAGGTCACATTCAGAGAGGAGTCCCGTCTGTGTACCAGCTACAGTA 618
|||||
190 rLysAsnTrpProAspHisAspValProSerSerIleAspProIleLeuG 207
|:::
619 TATGCTCTGGCCAGACCGTGGGTCTCCAGCAGTCTGACCAACATGCTCG 668
|||||
207 LuLeuIleTrpAspValArgCysTyrGlnGluAspAspSerValProIle 223
|||||
669 CCATGCTGGAGGAGCCGCTCGCTCCAGGGATCTGGCCCTGAACCCCTC 718
|||||
224 CysIleHisCysSerAlaGlyCysGlyArgThrGlyValIleCysAlaI 240
|||||
719 TGTGTCCACTGCAGTGGGGTGTGGGGAACAGCGGCTCTGTGTGCACC... 766
|||||
240 eValAspTyrThrTrpMetLeuLysAspGlyIleIleProGluAsnP 257
|:::
767 .GTGCATTATTGTGAGGACAGTCTCTGACCCAGATGATCCACCTGACT 815
|||||
257 heSerValPheSerLeuIleArgGluMetArgThrGlnArgProSerLeu 273
|||||
816 TCAGTCTCTTTGATGTGTCTTAAGATGAGGAACGCGGCTCGGCC 865
|||||
274 ValGlnThrGlnGluGlnTyrGluLeuValTyrAsnAlaValLeuGluLe 290
|||||
866 GTGCAGACAGAGGAGCAGTACAGGTTCCTGTACCACACGGTGGCTCAGAT 915
|||||
290 uPheLysArgGlnMetAspValIleArgAspLysHisSerGlyThrGluS 307
:|||||
916 GTTC.....TGCTCCACACTCCAGAAATGCCAGCCCCCCTACCAGA 956
|||||
307 exGlnAlaLysHisCysIlePro 314
::
957 ACATCAAGAGAAATGTGCCCA 979

```



```

alignment_block:
US-09-600-358A-4 x AAI58418      ..
Align seg 1/1 to: AAI58418 from: 1 to: 4798

    2 AspGlnArgGluIleLeuGlnLys..... 9
1682 GAACAGCAGGAGCTGTGGTGAAGGGATGATGAAACAACACTGATGAATGT 1731
10 .....PheLeuAspGluAlaGlnSerLysLysIle 20
1732 GGAGCCCAATCCATGCAGATATTGTTGGAACCTATAAGAGGAAAGATTG 1781
20 hrLysGluGlu.....PheAlaAsnGluPheLeuLysLeuLysArgGln 34
:: ||| ||||| ::::: ||| ::::: |||
1782 CTGATGAAGAAGACCTTTTCCTGGCGTGAAATTCACAGCATCCCGCGGGTG 1831
35 SerThrLysTyTLysAlaAspLysThrTyProThrThrValAlaGluAs 51
:: ||| ::::: ||| ::::: ||| ::::: |||
1832 TTCACGAG.....TTTCCTATAAAGAGAGCTCGAAA 1863
51 nAlaLysAsnIleLysLysAsnArgTyrllysAspIleLeuProTyAspt 68
:: ||| ::::: ||||| ::::: ||| ::::: |||
1864 GCCCTTTAACCCAGATAAAAAACCGTTATGTTGACATTTCTTCCTTAAT 1913
68 yrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSertyr 84
||| ::::: ||||| ||||| ::::: ||| ::::: |||
1914 ATAACCGTGTGAACCTCTGAGATAAACCAGGAGATGCAGGGTCAAAC 1963
85 IleAsnAlaAsnPheIleLysGlyValTyGlyProLysAlatyrIleAl 101
||| ::::: ||||| ||| ::::: ||| ::::: |||
1964 ATAAATGCCAGCTATTGATGTTTTCAAGAACCCAGGAATACATATGC 2013
101 aThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIlet 118
||||| ::::: ||||| ::::: ||||| ::::: |||||
2014 TGCACAAGGTCCCAGGGATGAACCTGTTGATCATTTCTCGAGGATGATTT 2063
118 rpGluTyTLysSerValleuIleIleValMetAlaCysMetGluTyrgluMet 134
||||| ::::: ||||| ::::: ||||| ::::: |||
2064 GGGAAACAGAAAGCCACAGTATTGTTCATGGTCACTCGATGTGAAGAAGGA 2113
135 GlyLysLysCysGluArgtyrTrpAlaGluProGlygluMetGlnLe 151
::::: ||||| ::::: ||||| ::::: |||||
2114 AACAGGAACAAGTGCAGCAATACTGGCCGTCAATGGAAGAGGGCACTCG 2163
151 uGluPheGlyPropheSerValSerCysGluAlaGluLysArgLysSera 168
||||| ::::: ||||| ::::: ||||| ::::: |||||
2164 GGCTTTTGGAGATGTTGTTGTAAGATCAACACAGCACAAAGATGTCAC 2213
168 spTyrlleIleargThrLeu.....LysValLysPheAsnSerGlu 181
||||| ::::: ||||| ::::: ||||| ::::: |||||
2214 ATTACATCATTCAGAAATTTGAACATTTGTAATAAAAAAAAAAAAAAGCA 2263
182 ThrArgThrIleTyrglnPheHisTyrllysAsnTrpProAspPhisAspv 198
||| ::::: ||||| ::::: ||||| ::::: |||||
2264 TGGAAGAGAGGTGACTCACATTCAGTTCCAGCTGGCCAGACCAGCGGG 2313
198 alProSerSerIleaspProIleLeuGluLeuIleTrpaspValArgCys 214
||||| ::::: ||||| ::::: ||||| ::::: |||||
2314 TGCCTTAGGATCTCTCACTTGTCTCTCAACTCAGAAGGAGAGTGAATGCC 2363
215 TyrGlnGluAspAspSerValProIleCysIleHIsCysSerAlaGlyCy 231
::::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
2364 TTCAGCAATTTCTCAGTGGGTCCCCATTGTGGTGCACTGCAGTGCTGGTGT 2413
231 sGlyArgThrGlyValIleCysAlaIleValAlaAspTyThrTrpMetLeuL 248
||||| ::::: ||||| ::::: ||||| ::::: |||||
2414 TGGGGCCACAGAACCTATATCGGAATT.....GATGCCA 2448
248 euLysAspGlyIlelleProGluAsn...PheSerValPheSerLeuille 263
::: ||||| ::::: ||||| ::::: ||||| ::::: |||||
2449 TGCTAGAAGGCTGGAAGCCGAGACAAGTGGATGTTTATGGTTATGTT 2498

```

[illegible]

520 AspSerSerAlaLeuGlyValTyr 527
|||::: ||| :::::
3324 AAGAGTACACCTCTACTCATTCAC 3347

seq_name: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:AAQ45353

seq_documentation_block:
ID AAQ45353 standard; DNA; 2790 BP.

AC AAQ45353;

XX 09-OCT-1994 (first entry)

XX Human protein-tyrosine-phosphatase-1D DNA.

XX Protein-tyrosine-phosphatase; enzyme; disease diagnosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 130..1911

XX FT /*tag= a

XX PN WO9408017-A.

XX PD 14-APR-1994.

XX PF 06-OCT-1993; 93WO-EP02728.

XX PR 06-OCT-1992; 92US-0956315.

XX PR 16-FEB-1993; 93US-0018129.

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PI Ullrich A, Vogel W;

XX DR WPI; 1994-135583/16.

XX DR P-PSDB; AAR52991.

XX PT New protein tyrosine phosphatase (PTP) protein, ptp-ID - are

XX PT useful for diagnosis and treatment of diseases associated with

XX PT abnormal PTP-ID levels

XX PS Claim 8; Page 64; 99pp; English.

XX CC This DNA may be expressed in a prokaryotic or eukaryotic host for

XX CC the production of a PTP-ID protein.

XX SQ Sequence 2790 BP; 861 A; 548 C; 669 G; 712 T; 0 other;

alignment_scores:
Quality: 497.00 Length: 441

Ratio: 1.972 Gaps: 15

Percent similarity: 57.143 Percent Identity: 31.293

alignment_block:
US-09-600-358A-4 x AAQ45353 ..

Align seg 1/1 to: AAQ45353 from: 1 to: 2790

14 AlaGlnSerLysLysIleThrLysGluGluPheAlaAsnGluPheLeuLy 30
|||::: ||| :::::

838 GCTGAGACCACAGATAAAGTCAACACAGGCTTTGGGAAGATTGAGAC 887
|||::: ||| :::::

30 sLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 47
|||::: ||| :::::

888 ACTACAACAACAGAGTGCAAACTT.....CTCTACAGCCGAA 925
|||::: ||| :::::

47 hrValAlaGluAsnAlaLysIleLysLysAsnArgTyrLysAspIle 63
|||::: ||| :::::

926 AAGAGGTCACAGCAAGAAAACAAAACAAATAGATATAAAACATC 975
|||::: ||| :::::

64 LeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspGl 80
|||::: ||| :::::
976 CTGCCCTTTGATCATACCAAGGTTGTCTCAGCATCTGTATCCCAATGA 1025
|||::: ||| :::::
80 uasp.....SerSerTyrIleAsnAlaAsnPheIleLysGlyValTyr.... 94
|||::: ||| :::::
1026 GCGTGTTCAGATTACATCAATGCAAAATATCATCATGCTCGATTTGAAA 1075
|||::: ||| :::::
95GlyProLysAlaTyrIleAlaThrGlnGly 104
|||::: ||| :::::
1076 CCAAGTGCACAAATTCAAAGCCCAAAAAGATTACATTGCCACACAGGC 1125
|||::: ||| :::::
105 ProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyrse 121
|||::: ||| :::::
1126 TGCGTCGAAAACACGCTGAATGACTTTTGGCGGATGCTGTCCCAAGAAA 1175
|||::: ||| :::::
121 rValLeuIleIleValMetAlaCysMetGluTyrGluMetGlyLysLysL 138
|||::: ||| :::::
1176 CTCCCGAGTGATTGTCATGACAAAGAAAGTGGAGAGAGGAAGAGTA 1225
|||::: ||| :::::
138 yscysGluArgTyrTrpAlaGluProGlyGluMetGlnLeu....GluPhe 153
|||::: ||| :::::
1226 AATGTGCTCAATACTGG.....CCTGATGAGTATGCTCTAAAAGAAATAT 1269
|||::: ||| :::::
154 GlyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrIl 170
|||::: ||| :::::
1270 GCGTCATCGCTTAGGAACGTCAAAAGAAAGCGCGCTCATGACTATAC 1319
|||::: ||| :::::
170 elleArgThrLeuLysVal.....LysPheAsnSerGluThra 183
|||::: ||| :::::
1320 GCTAAGAGAACTTAAACTTTCAAGGTTGGACAAAGGAATACGGAG...A 1366
|||::: ||| :::::
183 rgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPro 199
|||::: ||| :::::
1367 GAACGGTCGCGCAATACCACTTTCGGACCTGGCCGACACACGCGTGGCC 1416
|||::: ||| :::::
200 SerSerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGl 216
|||::: ||| :::::
1417 AGCGACCTGGGGCGTGTCTGGACTTCTCTGGAGGAGGTGCACCATGAAG 1466
|||::: ||| :::::
216 nGluAsp.....AspSerValProIleCysIleHisCysSerAlaGlyC 231
|||::: ||| :::::
1467 GGAGAGCATCATGGATGACGGCGCGTGTGTGCTGCTGCTGCTGCTGGA 1516
|||::: ||| :::::
231 ysglyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeu 247
|||::: ||| :::::
1517 TTGCCCGGACAGGAGCGTTC...ATTGTGATTGATATCTTATTGACATC 1563
|||::: ||| :::::
248 LeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuIleAr 264
|||::: ||| :::::
1564 ATCAGAGAGAAACGTGTGTGACTGCGATATTGACGTTCCCAAAACCATCCA 1613
|||::: ||| :::::
264 gGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrG 281
|||::: ||| :::::
1614 GATGGTGCAGTCTCAGAGGTCAGGGATGGTCCACAGACAGACAGTACC 1663
|||::: ||| :::::
281 luLeuValTyrAsnAlaVal.....LeuGluLeuPheLysArgGln 294
|||::: ||| :::::
1664 GATTATCATATATGCGGCTCCAGCATTTATATTGAACACACTACAGCGCAGG 1713
|||::: ||| :::::
295 MetAsp..... 296
|||::: ||| :::::
1714 ATTGAAGACAGCAGAAAACAAAGAGGAAAGGCACGAATATACAATAT 1763
|||::: ||| :::::
297ValIleArgAspLysHisSerGlyThrGluSerGlnAlaLysH 311
|||::: ||| :::::
1764 TAAGTATTCTCTAGCGGACACAGAGTGGATGAGATCAGAGCCCTCTCCCG 1813
|||::: ||| :::::
311 isCysIleProGlu..... 315
|||::: ||| :::::
1814 CTGTGACTCTCTGCCACCCTGTGCAGAAATGAGAGAGACAGTGTCTAGA 1863
|||::: ||| :::::
316LysAsnHisThrLeuGlnAlaAspSerTyrSer.....Pr 327
|||::: ||| :::::


```

87  aasnPhelIeLysGlyValTyGlyProLysAlaTyrlleAlaThrGlnG 104
|||||:||||| ||| |||:||||| |||:||||| |||:|||||
2847 CAACATATTGATGGCTACCAGACCAAGTCAATCATATTGCAACCAAG 2896
104  lyProLeuSerThrThrLeuLeuAspPheTrpArgMetIleTrpGluTyr 120
|||||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
2897 GTCCCGTTCATGAACACGTGTATGATTTCTGGAGGATGATTGGCAAGAA 2946
121  SerValLeuIleIleValMetAlaCysMetGluTyGluMetGlyLysLy 137
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
2947 CAATCTGCTGTCATTGCTGATGTTACAAATTTAGTTGAGTTGGCCGGT 2996
137  slyCysGluArgTyTrpAlaGluProGlyGluMetGlnLeuGluPheG 154
|||||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
2997 TAAATGCTATAAATATTGCGCTGATCACTGAAGTT.....TATG 3037
154  lyProPheSerValSerCysGluAlaGluLysArgLysSerAspTyrlle 170
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3038 GTGACTTCAAGTAACTGCTGTAAGAAATGGAACCACTTGCTGAATATGTA 3087
171  IleArg.....ThrLeuLysValLysPheAsnSerGluThrArgThrI1 185
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3088 GTTAGACATTCACCTCGAAGGAGGGGTACAAATGAATCCGTGAAGT 3137
185  eTyrlnPhelHisTyrlsAsnTrpProAspHisAspValProSerSerI 202
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3138 TAAACAGTTCCATTTACCGGCTGACCTGACCATTGAGTGCCCTACCATG 3187
202  leAspProIleLeuGluIleTrpAspValArgCysTyrlnGlnLysPhe 218
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3188 CPACAGGCTGCTTTCCTTTATCCGGCGAGTCAAGTTATCAACCCCTCC 3237
219  AspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgThrGl 235
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3238 AGTGTGCGCCCATCGTGTGATTCAGTGTGCTGTGCTGGACGACATGG 3287
235  yValIleCysAlaIleValAspTyThrTrpMetLeuLeuLysAspGlyI 252
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3288 CTGTACATTTGTGATGACATCATGCTAGACATGGCTGAAGAGAGGGTG 3337
252  leIleProGluAsnPheSerValPheSerLeuIleArgGluMetArgThr 268
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3338 TTGTT.....GATATTTCAATTTGTCAAAAGCCCTTAAGATCT 3375
269  GlnArgProSerLeuValGlnThrGlnGluIntyrGluLeuValTyraS 285
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3376 CGCGCTATTATATGTCAGACAGAGGAACAGTACATTTTATTTCATGA 3425
285  nAlaValLeuGluLeu..... 290
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3426 TGCCATTTTAGAAGCTGCTTATGTGGAGAAACTGCCATACCTGCTGTG 3475
291  ...PheLys...ArgGlnMetaspValIleArg..... 299
||||| |||:||||| |||:||||| |||:||||| |||:|||||
3476 AATTTAAAGCTGCATATTTGATGATTAGATAAGTACTCCAGACTAAC 3525
300  AspLysHisSerGlyThrGluSerGlnAlaLysHis..... 311
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3526 TCTTCACATCTCAAGGATGAATTCAGACTCTGAATTCAGTACCCCTCG 3575
312  ..... 319
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3576 ACTACAAGCTGAAGACTGCAGTATAGCTGCTGCCA...AGGAACCATG 3622
319  hrLeuGln.....AlaAspSerTyrlsProAsn.....LeuProLys 331
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3623 ACAAGAACCCTTTCATGAGACATGCTGCCACCTGCACAGATGCTGCTTTT 3672
332  SerThrThrLysAlaLysMetMetAsnGlnGlnArgThrLysMet... 347
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3673 TTAATTACAATTTGAGGAGAGAGTACTACATCAATGCTGCTCTTAT 3722
348  ....GluIleLysGluSerSerPheAspPheArgThrSerGluIleS 363

```

```

|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3723 GGACAGCTACAGGCAACCAAGCTGCTTTC..... 3750
363  erAlaLysGluGluLeuValLeuHisProAlaLysSerSerThrSerPhe 379
|||||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3751 .....ATCGTCACACAAATACCCCTCTGCCAAACACTGTAAAA 3786
380  AspPheLeuGluLeuAsnTyrlsSerPhe..... 388
|||||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3787 GACTTCTGGAGATTAGTGTATGATTATGCTGTACCTCCATTGTGATGTT 3836
389  .....AspLysAsnAlaAspThrThrMetLysTrpGlnThrLysA 402
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3837 AAACGAAGTCGACTTGTCCAGGCTGCCCTCAGTACTGGCCAGAGGAAG 3886
402  laPhe.....ProIleValGlyGluProLeuGlnLysHisGln 414
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3887 GGATCTACGATATGGCCCATCCAAAGTGAATGATGTCTTTGTTCAATG 3936
415  SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3937 GACTGTGATGTATCAACCGGATTTTAGGATATGCAATCTAACAGACC 3986
431  oValAsnAla.....AlaGlyArgTyrlsP 439
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
3987 ACAGGAAGTATTATCTGATGGTGCAACAGTTTCAGTACCTAGGATGGCTT 4036
439  heAsnSerLysValProIleThrArgThrLysSerThrProPheGluLeu 455
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
4037 CTCATCGAGAGTGCCT.....GGATCCAAAGGTCATCTTTGAACTG 4080
456  ileGlnGlnArgGluThrLysGluValAspSerLysGlu 468
|||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
4081 ATACTTCAGGTGAAAGTGGCAGGAGGAATGCCAAGGAG 4119

```

OM of: US-09-600-358A-4 to: GenEmbl.* out_format : pfs
Date: May 26, 2002 2:12 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09600358/runat_24052002_101554_27006/app_query.fasta_1.761
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DEEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09600358_@CNG1_1_11307
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-600-358A-4
Query length: 692
Database: GenEmbl.*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 3968.550000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_pat:AX020263	+ 3615.00	5088.29	4.7e-275	2356	AX020263 Sequence 3 from Paten
gb_pr:AF001847	+ 3611.00	5082.64	9.6e-275	2356	AF001847 Homo sapiens lymphoid
gb_pat:AX020261	+ 3573.00	5026.17	1.3e-271	3058	AX020261 Sequence 1 from Paten
gb_pr:AF001846	+ 3569.00	5020.51	2.8e-271	3058	AF001846 Homo sapiens lymphoid
gb_pat:AR149916	+ 3543.50	4982.80	3.5e-269	3580	AR149916 Sequence 1 from Paten
gb_pr:AF150732	+ 2977.50	4185.30	9.2e-225	2877	AF150732 Homo sapiens protein
gb_to:MUSPROTYPH	+ 2356.50	3308.29	6.5e-176	2734	M90388 Mouse protein tyrosine
gb_pr:AL137856	- 1490.50	2042.33	2.1e-105	143618	AL137856 Human DNA sequence
gb_pr:HUMPTPGL	+ 1157.00	1612.48	1.9e-81	2938	DI3380 Human mRNA for protein-
gb_sts:G18337	+ 1157.00	1611.71	2.1e-81	3160	G18337 sWSS1985 Eric D. Green
gb_pr:HUMPTPPEST	+ 1157.00	1611.71	2.1e-81	3160	M93425 Human protein tyrosine
gb_to:MMXPTPPES	+ 1130.50	1576.03	2.0e-79	2676	X86781 M.musculus mtp-pest ge
gb_pat:AX306312	+ 1104.50	1538.13	2.6e-77	2983	AX306312 Sequence 1063 from Pa
gb_to:MMP19PTP	+ 1058.00	1479.00	5.1e-74	1608	D38072 Rat mRNA for protein ty
gb_to:RATRKPTP	+ 767.00	1066.48	4.8e-51	1817	BC017785 Homo sapiens, clone M
gb_pr:BC017785	+ 748.00	1037.61	2.0e-49	2198	U69673 Rattus norvegicus prote
gb_to:RNU69673	+ 748.00	1037.47	2.0e-49	2226	AR095603 Sequence 1 from Paten
gb_pat:AR095603	+ 736.50	1025.22	9.6e-49	1529	AR154418 Sequence 1 from Paten
gb_pat:BC008512	+ 735.50	1023.04	1.3e-48	1643	BC008512 Mus musculus, Similan
gb_to:MMPTPHSC	+ 733.50	1020.98	1.7e-48	1529	U49853 Mus musculus protein ty
gb_to:MMU35124	+ 733.50	1020.78	1.7e-48	1559	U35124 Mus musculus nonrecept
gb_to:MMU52523	+ 727.50	1012.24	5.1e-48	1568	U52523 Mus musculus protein ty
gb_pr:HSBDP1	+ 710.00	981.29	2.7e-46	2810	X79568 H.sapiens BDP1 mRNA for
gb_to:MUSLY5AA	+ 521.00	708.06	4.5e-31	5007	M92933 Mouse lymphocyte common
gb_pr:HSICAR	+ 516.50	703.28	8.2e-31	4315	U00638 Human mRNA for leukocy
gb_pr:AF173857	+ 515.50	701.20	1.1e-30	4597	Y00062 Human mRNA for T200 leu
gb_to:AF173857	+ 515.50	700.47	1.2e-30	4920	AF173857 Xenopus laevis recept
gb_to:MUSC57BL6A	+ 513.00	710.40	3.3e-31	1391	DA9393 Mouse mRNA for protein
gb_to:DB3203	+ 513.00	698.60	1.5e-30	4208	D83203 Mus musculus mRNA, one
gb_to:MUSBTP	+ 513.00	698.07	1.6e-30	4426	D45212 Mouse mRNA for receptio
gb_to:DB3204	+ 513.00	697.67	1.7e-30	4596	D83204 Mus musculus mRNA, one
gb_to:MUSLY5A	+ 512.00	696.12	2.1e-30	4655	M14342 Mouse 200 kD T-cell leu
gb_to:MUSLY5B	+ 512.00	694.97	2.4e-30	5186	M14343 Mouse Ly-5 (leucocyte-d
gb_pat:A84503	+ 509.00	698.24	1.6e-30	2562	A84503 Sequence 118 from Paten
gb_pr:HUMPTPES	+ 508.00	690.24	4.4e-30	4756	D37781 Human mRNA for protein-
gb_pat:AR109903	+ 508.00	689.46	4.9e-30	5117	AR109903 Sequence 1 from Paten
gb_pr:HSU10886	+ 508.00	689.46	4.9e-30	5117	U10886 Human density enhanced
gb_to:RNU40790	+ 508.00	687.22	6.5e-30	6314	U40790 Rattus norvegicus vascu
gb_to:RATLCA1	+ 502.50	685.61	7.9e-30	3541	M10072 Rat leukocyte common ad

gb_ro:RNLICAR + 502.50 684.24 9.5e-30 4027 ! Y00065 Rat mRNA for leucocy
gb_ov:AB031424 + 500.50 680.25 1.6e-29 4490 ! AB031424 Cyprinus carpio mR
gb_ro:AY026862 + 498.00 677.08 2.4e-29 4338 ! AY026862 Mus musculus brain
gb_to:AY026861 + 498.00 677.02 2.4e-29 4365 ! AY026861 Mus musculus brain
gb_to:AF162856 + 498.00 676.55 2.5e-29 4562 ! AF162856 Mus musculus recep

seq_name: gb_pat:AX020263

seq_documentation_block:
LOCUS AX020263 2356 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 3 from Patent WO9936548.
ACCESSION AX020263
VERSION AX020263.1 GI:10044044
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2356)
AUTHORS Roifman,C.M.
TITLE Human lymphoid protein tyrosine phosphatases
JOURNAL HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)
FEATURES
source
1..2356
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2120
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07769.1"
/db_xref="GI:10044045"
/translation="MQQREILQKFLDEAQSKTKKEEFANEFLLKRSQTKYKADKTY
PTVAENAKNIKRNKRDILPDYISRVESLTSDESSYINANFINGVVGPKAYIAT
QGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKKCKERYWAEPPGEMLEFGPFVSVC
EAEKRSKDIYIRTLKVKFNSETIYQHYKNPDRHDPSSIDPILLELWDVRCYQED
DSVPICITGSCAGCGRTVCAIVDYTWMLIKDGIIPENSVFSLIREMRTORSLVQT
QOEYELVYNAVLRLFKRQMDVIRDKHSGTSEQAQKHCIPEKNHTLQADYSNLPKSTT
KAAKMNQQTQKEIKESSEDFRTSEISAKELVLHPAKSTSFDFLELNYSEFNA
DTMKWQTKAFPIVLEPLQHQSLDGLSFLFEGCSKSPVNAAGRFTNSKVPITRTKS
TPFELIQQRETRKEDSKENFSYLPESQHDSCFVEMQAQKVMHVSSAELNYSLPYDSKH
QIRNASNHRHDSKALGVSYIPLVENPYSPWPPSGTSKMSLDLPEKQDGVFPFSK
LLPTSTSLFYSYNSHSSLSLNSPTNISLLNQESAVLATAPIDDEIPPLPVRPPE
SFTVVEAGEFSPNVKPSLSAVKVKIGTSLWGGTSEPKKFDSDVILRPSKSVKLRS
PKSGKNFSL"
BASE COUNT 750 a 466 c 466 g 674 t
ORIGIN

alignment_scores:
Quality: 3615.00 Length: 692
Ratio: 5.224 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-600-358A-4 x AX020263

Align seg 1/1 to: AX020263 from: 1 to: 2356

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
|||||
42 ATGGACCAAGAGAAATTCGAGAAATTCCTGGATGATGAGCCCAAGCAA 91
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92 GAAATATTACTAAAGAGAGATTTCGCAATGAATTTCTGAAGCTGAAAGGC 141
34 InSerThrLysTyrlsAlaAspLysThrTyProThrThrValaAlaGlu 50
|||||
142 AATCTACCAAGTACAGGACAGCAAAACCTATCCTACAACTGTGCTGAG 191
51 AsnAlaLysAsnIleLysLysAsnArgTyrlsAspIleLeuProTyra 67
|||||

192 AATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGGCCCTATGA 241
67 pTyrSerArgValGluLeuSerLeuileThrSerAspGluAspSerSert 84
|||||
242 TTATAGCGGGTAGAAGTATCCCTGATGAACCTCTGATGAGGATTCAGCT 291
84 yrtleAsnAlaAsnPheileLysGlyValTyrGlyProLysAlaTyrile 100
|||||
292 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCAAGGCTATATT 341
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpargMetII 117
|||||
342 GCCACCCAGGGCTCTTATCTACAACCCCTCTGGACTTCGGAGGATGAT 391
117 eTrrpGluTyrSerValLeuileleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTGCTCTTATCATTTGTATNGGCATGTCATGGAGTATGAAA 441
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAAAAGAAAGTGTGAGCGCTACTGGGCTGAGCCGAGAGATGCAG 491
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysargLysSe 167
|||||
492 CTGGAAATTTGGCCCTTTCTCTGTATCTCTGTAAGCTGAAAAAGGAAATC 541
167 rAspTyrileileArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TGATTATATATCAGGACTCTAAAGTTTAAGTTCAATAGTAGAACTCGAA 591
184 hrileTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTACCAGTTTCATTACAAGAAATGGCCAGACCATGATGATACCTTCA 641
201 SerileaspProileLeuGluLeuiletrPaspValargCysTyrGlnGl 217
|||||
642 TCTATAGACCCCTATTCTTGAGCTCATCTGGGATGTACGTTGTTTACCAAGA 691
217 uAspAspSerValProileCysileHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTCCCATATATGCAATCTACTGCAGTGTGGCTGTGGGAAGA 741
234 hrGlyValileCysAlaileValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTGTCGTATTGTTGATTATACATGGATGTTGCTTAAAGAT 791
251 GlyleileProGluAsnPheSerValPheSerLeuileargGluMetAr 267
|||||
792 GGGATAATTCCTGAGAACTTCAGTGTGTTTCAGTGTGATCCGGGAAATGCG 841
267 qThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCTTCATTAGTTTCAACGACGAGAAACAATATGAAGTGTCT 891
284 yrAsnAlaValLeuGluLeuPheLysargGlnMetAspValileArgAsp 300
|||||
892 ACAATGCTGTATTAGAAGTATTTAAGACAGATGGATGTTATCAGAGAT 941
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysileProGluLysAs 317
|||||
942 AAACATTTCTGGACAGAGAGCTCAAGCAAGCAATTTGATTCCTTGAGAAAA 991
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCACGACACTCTTATCTCCTTAATTTACCAAAAGTACCA 1041
334 hrLysAlaLysMetMetAsnGlnGlnArgThrLysMetGluileLys 350
|||||
1042 CAAAAGCAGCAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1091
351 GluSerSerPheAspPheArgThrSerGluileSerAlaLysGluGl 367
|||||
1092 GAATCTTCTTCTGACTTTAGGACTTCTGAAATAAGTGCAAAAGGA 1141

367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTGGCACCCCTGCTAAATCAAGCACTTCTTTTGACTTCTTGAGC 1191
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAAATTTACAGTTTTTGACAAAATGCTGACAAACCATGAATGGCAGACA 1241
401 LysAlaPheProileValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAATAGTTGGGAGCCCTTTCAGAAGCATCAAAAGTTTGA 1291
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTTGTGAGGGATGTTCTAATTTCTAAACCTGTAATG 1341
434 laAlaGlyArgTyrPheAsnSerLysValProileThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTTAAATTCAAAGGTGCGCAATAACACGAGCAATCA 1391
451 ThrProPheGluLeuileGlnGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCCCTTTTGAATGTACAGCAGAGAGAAACCAAGGAGGTGGACAGCAA 1441
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTTCTTATTTGGAATCTCAACCCACATGATTTCTTTTGTAG 1491
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCGAGGCTCAAAAAGTAAATGCATGTTTCTTCAGCAGAACTGAATTAT 1541
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCCTGCGCATATGACTCTAAACACCAAAATACGTAATGCCTCTAATGTAAA 1591
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTGTCTTGTGTATATTTCTATACATACCTTTAGTGG 1641
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGCGCTCCAAGTGTGTACCACTTCTAAGATG 1691
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerle 567
|||||
1692 TCTCTTGATTTTACCTGAGAAGCAAGATGGAACGTGTTTTCTTCTCTCT 1741
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCAACATCCTCTACATCCCTTCTTCTTATTACAATTCACATAGTT 1791
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCAGTGAATTTCTCCAACCAATATTTCTCCTACATTTTGAACCCAGAG 1841
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCACTGCTCCAAGGATAGATGAATTCCTCCCTCC 1891
617 oLeuProValargThrProGluSerPheileValValGluGluAlaGlyG 634
|||||
1892 ACTTCTGTACGGACACCTGAATCATTTATTGTGTTGAGGAAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAATGTTCCCAAAATCCTTATCTCAGCTGTGAAGGTAATA 1991
651 IleGlyThrSerLeuGluTyrGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAAACATCAGTGGAAATGGGGTGGAACTCTGACCAAGAAATTTGA 2041

667 pAspServAlileuArgProSerLysSerValLysLeuArgSerProL 684
 |||||
 2042 TGACTCTGTGATCTAGACCAAGCAGAGTGTAAACTCCGAGTCCTA 2091
 684 ySserGlyLysAsnPheSerTrpLeu 692
 2092 AATCAGGTAAATTTCTCTGTGGCTT 2117

seq_name: gb_pr:AF001847

seq_documentation_block:
 LOCUS AF001847 2356 bp mRNA linear PRI 01-OCT-2001
 DEFINITION Homo sapiens lymphoid phosphatase Lyp2 mRNA, complete cds.

ACCESSION AF001847
 VERSION AF001847.1 GI:4100633

KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2356)
 REFERENCE Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. and Roifman, C.M.
 AUTHORS Cloning and characterization of a lymphoid-specific, inducible
 TITLE human protein tyrosine phosphatase, Lyp
 JOURNAL Blood 93 (6), 2013-2024 (1999)

MEDLINE 99168989
 PUBMED 10068674

REFERENCE 2 (bases 1 to 2356)
 Roifman, C.M.
 AUTHORS Direct Submission
 TITLE Submitted (30-APR-1997) Immunology and Allergy, The Hospital For
 JOURNAL Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,
 Canada

FEATURES Location/Qualifiers

1..2356
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 42..2120
 /note="splice variant of lymphoid phosphatase Lyp1"
 /codon_start=1
 /product="lymphoid phosphatase Lyp2"
 /protein_id="AAD00905.1"
 /db_xref="GI:4100634"
 /translation="MDQREILQKFLDEAQSKEITKEEFANFELKLRQSTKYKADKTY
 PTVVAENAKIKKNRYKIDILPYDSRVLSLISDESSINANFIKGVGPYAIAT
 QGPLSTLLDFRWINEYSVLIVMACMEYEMGRKKERWAEPGEMLQLEFGPPSVSC
 EAERKSDYIIRTLKVFNSERTIYQFYHKNWPDHPDPSIDPLELIWDVRCYQED
 DSVPICTHCSAGCGRTGVCIAVDYTWLLKDGITIPENFSVFLIREMRTQPSLVDT
 QEOVELYNVAVLELFKROMDVIRDKHSGTESQAKHCIPKRNHTLOADSYSPNLPKSTT
 KAAKMMNQRTKMEIKSSSFDPRTSEISAKEELVLPKASSTSFDFLELYNPFDKNA
 DTNNKWTQAPPIVGEPLQKHSLDLSLLEFGCSNKPYNNAAGRYFNKVPITRTKS
 TPFLIQORETKEDVDSKENSYLEQSPHDSCFVEMQAKQVHVHSSAELNYSLPYDSKH
 QIRNASNVKHHDSALGVYIYIPLENPFYSSWPPSGTSKMSLDLPEKQDGVFPSS
 LLPTSSLSFYSYNSHDSLSINSPNISLNQESAVLATAPRIDDEIPPLPVPRTPE
 SFIVVEEAGFSPNPKVSLSSAVKVIKTSLEWGTSEPKKFDSDVILRPSSKVKLRS
 PKSGKNFSWL"

BASE COUNT 750 a 466 c 466 g 674 t
 ORIGIN

alignment_scores:
 Quality: 3611.00 Length: 692
 Ratio: 5.218 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.855

alignment_block:
 US-09-600-358A-4 x AF001847 ..

Align seg 1/1 to: AF001847 from: 1 to: 2356

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
 |||||
 42 ATGGACCAAGAGAAATTCGACAGAGTTCTCTGGATGAGGCCCAAGCAA 91

17 sLysIleThrLysGluLupPheAlaAsnGluPheLeuLysLeuLysArgG 34
 |||||
 92 GAAATTTACTAAGAGAGAGTTTGGCCAATGAATTTCTGAAGCTGAAAGGC 141
 34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
 |||||
 142 AATCTACCAAGTACAAAGGCAGACAAAACCTATCTACAACTGTGGCTGAG 191
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 |||||
 192 AATGCCAAGAATATCAAGAAAAACACATATAAGGATATTTGCCCTATGA 241
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerT 84
 |||||
 242 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGTATGAGGATTCAGCT 291
 84 yTleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 |||||
 292 ACATCAATGCCAACTTCATTAAGGGAGTTTATGGACCCCAAGGCTTATAT 341
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
 |||||
 342 GCCACCCAGGGCTCTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 391
 117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGlu 134
 |||||
 392 TTGGGAATATAGTGTCTTATCATTTATGGCATGCATGGAGTATGAAA 441
 134 eTGLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 |||||
 442 TGGGAAAGAAAAGTGTAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 491
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
 |||||
 492 CTGGAATTTGGCCCTTCTCTGTATCTGTGAAGCTGAAAGGAAAGAAATC 541
 167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
 |||||
 542 TGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACATCGAA 591
 184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProser 200
 |||||
 592 CTATCTACCAGTTTCATTAACAAGATTGGCCAGACCATCATCTACCTTCA 641
 201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
 |||||
 642 TCTATAGACCTTATCTTGGAGCTCATCTGGGATGTACGTTGTACCAAGA 691
 217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
 |||||
 692 GGATGACAGTGTTCATCATATGATTCACATGCAGTGTGGCTGTGGAAGA 741
 234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
 |||||
 742 CTGGGTGTTATTTGCTATTGTTGATTATACATGGATGTTGCTAAAAGAT 791
 251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetar 267
 |||||
 792 GGGATAAATTCCTGAGAACTTCAGTGTGTTTTCAGTTTTCATCCGGGAAATGCG 841
 267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
 |||||
 842 GACACAGAGGCCCTTATAGTTAGTCAACACGAGGAGCAACAATAGACTGGTCT 891
 284 yTAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
 |||||
 892 ACAATGCTGTATTAGAACTATTATTAGACACACATGGATGTTATCAGAT 941
 301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
 |||||
 942 AAACATTTCTGGAACAGAGAGTCAAGCAAGCAATGTTGTTCTTCTGAGAAAAA 991

```
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThr 334
|||||
992 TCACACTCTCCAAGCAGACTCTATTCTCTAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluLeLys 350
|||||
1042 CAAAAGCAGCAAAAATGATGAACCAACAAGGAGCAAAAATGGAATCAAA 1091
|||||
351 GluSerSerPheAspPheArgThrSerGluLeuSerAlaLysGluG1 367
|||||
1092 GAATCTTCTCTCTTGAATTTAGGACTTCTGAATAAGTGCAAAAAGAGA 1141
|||||
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTCACCCCTGCTAAATCAAGCACTTCTTTTGACTTCTCGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTyrGlnThr 400
|||||
1192 TAAATTACAGTTTTGACAAAAATGCTGACACAACCATGAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCAATAGTTGGGAGCCCTTCAGAAAGCATCAAAAGTTTGGG 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTTGTGTAGGGATGTTCTAATTTCTAAACCTGTAATG 1341
|||||
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAATTCAAAGGTGGCAATTAACACGAGCAACCAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCTTTTGAATTTGATACAGCAGAGAGAACCAAGGAGGTGGACAGAA 1441
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTTCTATTGGAATCTCAACACATGATTCGTGTTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCGAGGCTCAAAAAGTATGATGATGTTCTTCAGCAGAACTGAATAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACATGCCATATGACTCTAAACACCAATAGTAAATGCCCTCTAATGAAA 1591
|||||
517 shHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTCTCTTGGTGTATATTTCTATACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTCTATTTTTCATCATGCGCTCCCAAGTGTACCACTTCTAAGATG 1691
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTTACCTGAGACAGATGGAAGTCTTTTCTCTCTCTCT 1741
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCACATCTCTACATCCCTCTTCTCTTATTACAATTCACATGATT 1791
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTTCTCAACCAATATTTCTCTCACTATTGAACCGAGG 1841
|||||
601 SerAlaValLeuAlaThrAlaProArgTleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCACTGCTCCAGGATAGATGATGAATCCCCCTCC 1891
|||||
617 oLeuProValArgThrProGluSerPheIleValIleGluAlaGlyG 634
|||||
```

```
|||||
1892 ACTTCTCTGTACGACACCTGAATCATTTATTGTGGTTGAGGAAGCTGGAG 1941
|||||
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAAATGTCCAAATCTCTATCTCAGCTGTGAAGGTAAAA 1991
|||||
651 IleGlyThrSerLeuGluTyrGlyGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAACATCACTGGAATGGGTGGACATCTGAACCAAGAATTTGA 2041
|||||
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTGATCTTAGACCAAGCAGAGTGTAATAACTCCGAAGTCCTA 2091
|||||
684 ySSerGlyLysAsnPheSerTrpLeu 692
|||||
2092 AATCAGGTAAAAATTTCTCTTGGCTT 2117
|||||

seq_name: gb_pat:AX020261

seq_documentation_block:
LOCUS AX020261 3058 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9936548.
ACCESSION AX020261
VERSION AX020261.1 GI:10044042
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3058)
AUTHORS Roifman,C.M.
TITLE Human lymphoid protein tyrosine phosphatases
JOURNAL Patent: WO 9936548-A 1 22-JUL-1999;
HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)
FEATURES
Location/Qualifiers
source
1..3058
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2468
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07768.1"
/db_xref="GI:10044043"
/translation="MDQELQKFLDEAQSKKITKEEFANFLKLKROSTKYKADKTY
PTTVAENAKNKKRYKDLIPYDSRVLSLTDESSYINANFIKGVYGPKAVIAT
EAKRKSIIIRTLKVFNSERTIYQHYKNPDHVPSSIDPLELIDVRCVQED
DSVPICHSAGCGRTGIVCAIVDYTWMLKDGIIIPENFSVFLIREMTRQPSLVQT
QEOVELVYNAVLELFRQMDVIRDKHSGESQAKHCIEKHNHTLQADSYSPNLPKSTT
KAAKMNQORTKMEIKESSDFDFTSEISAKEELVLPKSKSTSFDFLELNYSPDKNA
DTTMKWTQAFPIVGEPLQKQSLDGLSGILLFEGCSNKPVRNAGRFNSKVPITRTKS
TFELIQRETKVEDSKENFSYLESQPHDSCFVEMQAOKVMHVSAELNYSILPDSKH
QIRNASNYKHHDSALGVSYIPLVNFYFSWPPSGTSSKMSLDLPKQDGTVPSS
LIPTSTSLFSYINSHSLNSPTNISSLNQESAVLATAPRIDEPPLPVPRTPE
SFIVVEAGEFSPNPVKLSAVKVIKIGTSLSEWGGTSEPKKFDSDVILRPKSKVLRS
PKSELHODRSSPPPLPERTLESFFLADEDCMQAOSIETYSYPTDWNSTSSKQTL
KTPGKSTFRSKSLKILRNKKKSCICNCPNKPAPESVQSNSSSFLNFGFANRFSKPKG
PRNPPTTWN!"
BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

alignment_scores:
Quality: 3573.00 Length: 685
Ratio: 5.216 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-600-358A-4 x AX020261 ..
Align seg 1/1 to: AX020261 from: 1 to: 3058
```


1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
|||||
42 ATGGACCAAGAGAAATTCGAGAGATTCCTGGATGAGGCCCAAGACAA 91
|||||
17 slyIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
92 GAAATTTACTAAAGAGAGAGTTTGGCAATGAATTTCTGAAGCTGAAAGGC 141
|||||
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 AATCTACCAAGTACAAAGGAGACAAACCTATCCTACAACTGTGGCTGAG 191
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 ANTGCCAAGAATATCAAGAAACAGATATAAGGATATTTGCCCTATGA 241
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242 TTATAGCCGGGTAGAACTATCCCTGATACCTCTGATGAGGATTCAGCT 291
|||||
84 yIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACTTCATTAAGGAGTTTATGGACCCCAAGGCTTATATT 341
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetIl 117
|||||
342 GCCACCCAGGGTCCCTTATCTACACCCCTCTGGACTTCGGAGATGAT 391
|||||
117 eTrrGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTGTCTTATCATTTGTTATGGCATGCATGGAGTATGAA 441
|||||
134 eGlyLysLysCysGluArgTyrTrrAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAGAAAGAAAGTGTGAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 491
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492 CTGGAAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAGAAATC 541
|||||
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TCAATATATATATACAGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAA 591
|||||
184 hrIleTyrGlnPheHisTyrLysAsnTrrProAspHisAspValProser 200
|||||
592 CTATCTACCAAGTTTCATTACAAAGAATTGGCCAGACCATGATGACCTTCA 641
|||||
201 SerIleAspProIleLeuGluLeuIleTrrAspValArgCysTyrGlnGl 217
|||||
642 TCTATAGACCCCTATTTCTTGAGCTCATCTGGGATGACGTGTACCAAGA 691
|||||
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTTCACATATGATTCACCTGCAGTGTGCTGTGGAAGA 741
|||||
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
742 CTGGTGTATTATTTGTGCTATTGTGATTATATACATGGATGTTGTAAGAAT 791
|||||
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGATAAATCCCGAAGCTTCAGTGTTCAGTTTTCAGTTTGGGGAATGCG 841
|||||
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCCCTTCATTAGTTCAACACGACGAGCAACAATATGAACGTGCT 891
|||||
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACAATGCTGTATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGAT 941
|||||

301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTTCTGGAACACAGAGACTCAAGCAAGCAATGTATTCTCTGAGAAAA 991
|||||
317 nhIsthrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAAGCAGACTTATTCTCTTAATTTACCAAAAAGTACCA 1041
|||||
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAAATGGAATCAAA 1091
|||||
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG 367
|||||
1092 GAATCTTCTCTCTTTCACCTTAGGACTTCTGAATTAAGTGCAAAAGAAGA 1141
|||||
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTGCACCCCTGCTAAATCAAGCACTTCTTTTGTACTTTCTGGAGC 1191
|||||
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrrpGlnThr 400
|||||
1192 TAAATTTACAGTTTTTGACAAAAATGCTGACACAAACCATGAAATGGCAGACA 1241
|||||
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AAGGCATTTCCAAATAGTTGGGAGGCTCTTCAGAGCATCAAGTTTGA 1291
|||||
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTTGTGTTGAGGATGTTCTAATCTAAACCTGTAATG 1341
|||||
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTAATCAAAAGGTGCCAATAACACGACCAATCA 1391
|||||
451 ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
|||||
1392 ACTCTCTTTTGAATTTGATCAGCAGAGAGAGAAACCAAGGAGGTGGACAGCAA 1441
|||||
467 sGluAsnPheSerTyrLeuGluSerGlnProHisaspSerCysPheValG 484
|||||
1442 GGAACAACTTTTCTTATTTGGAATCTCAACACATGATTTCTGTTGTAG 1491
|||||
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCAGGCTCAAAAAGTAATGCATGTTTCTTCAGCAGAACTGAATAT 1541
|||||
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1542 TCACGCCCATATGACTCTAAACACCAATACGTAATGCTCTAATGTAAA 1591
|||||
517 shHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTCTAGTGTCTTGGTGTATATTCTTACATACCTTTAGTGG 1641
|||||
534 luAsnProTyrPheSerSerTrrProProSerGlyThrSerSerLysMet 550
|||||
1642 AAAATCCTTATTTTTCATCATGGCCTCCCAAGTGGTACCAGTTCTAAGATG 1691
|||||
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTACCTGAGAGCAAGATGGAAGTGTTCCTTCTCTCTCT 1741
|||||
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTGCCAAACATCCTCTACATCCCTCTCTCTTATTACAAATTCACATAGTT 1791
|||||
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTCCTCAACCAATATTTCTCCTACTATTGAACGAGGAG 1841
|||||
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProPr 617
|||||

|||||
1842 TCAGCTGTACTAGCAACTGTCTCCAGGATAGATGATGAAATCCCCCTCC 1891
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
1892 ACTTCCCTGTACGGACACTGAATCATTTATGTGTGGTGAAGAGCTGGAG 1941
634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
1942 AATTCTCACCAAAATGTTCCCAAAATCCTTATCTCAGCTGTGAAGGTAAAA 1991
651 lIedGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
1992 ATTGGAACATCACATGGAATGGGTGGAACATCTCGAACCAAGAAATTTGA 2041
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
2042 TGACTCTGTGATACTTGAACCAAGAGGTGAATAACTCGGAAGTCCCTA 2091
684 ySSer 685
|||||
2092 AATCA 2096

seq_name: gb_pr:AF001846

seq_documentation_block:

LOCUS AF001846 3058 bp mRNA linear PRI 01-OCT-2001
DEFINITION Homo sapiens lymphoid phosphatase LyP1 mRNA, complete cds.

ACCESSION AF001846

VERSION AF001846.1 GI:4100631

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3058)

Cohen,S., Dadi,H., Shaoul,E., Sharfe,N. and Roifman,C.M.

Cloning and characterization of a lymphoid-specific, inducible

human protein tyrosine phosphatase, Lyp

Blood 93 (6), 2013-2024 (1999)

99168989

10068674

2 (bases 1 to 3058)

Roifman,C.M.

Direct Submission

Submitted (30-APR-1997) Immunology and Allergy, The Hospital For

Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,

Canada

Location/Qualifiers

1..3058

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="thymocytes"

42..2468

/codon_start=1

/product="lymphoid phosphatase LyP1"

/protein_id="AAD0904.1"

/db_xref="GI:4100632"

/translation="MDQREILQKFLDEAOSKKITKEEFANFLKLRQSTKYKADTY

PTVAENANIKNRKYDILPYDYSRVELSLTSDSDSYINANFIKGYGPKAYIAT

QQLPDLDFRMWTEYISVLIIYMACMEYEMGKKKERYWAEPEGMELEFGPFYSVC

EAEKRSYIIIRTLKVFNSERTIYQFYKNWPDHVPSSIDPILIELIWDVRCYQED

DSVPICISACGGRTGVTCAIVDTWMLLKQIIPENSVFSLIREMTORPSIVOT

QOQELVYNAVLELFRKQNDVIRDKHSGTFSQAKHCIPENKHTLOADSYSPNLKSTT

KAAMNQWRTWEIKESSDFPRTSEISAKBELVLHPAKSTSFDFLELNSFPKNA

DTMKWQTKAFPIVGLPQKHSLDGLSLLFEGCSNSKPVNAAGFYNSKVPITRTKS

TFPELQQERTKEVSKNFYSLEQPHDSCFVEMQAKVHVSSAELNYSLPYDSKH

QIRNASQVHRHSDSLGVSYIPLVENPYSFSSPGTSKMSLDLPEKQDGTVPFS

LIPTSLVFNYSNHDLSLSPNISLLNOSAVLATAPRIDDEIPPLPVPTPE

STIVVEAGEFSFNPYKPSLSAAVKVKGTSLEWGGTSEPKFDDSVILRPSKVLRS

PKSELHQDRSSPPPLPERTLESFLEADEDCHQAOISIEYTSYPTDMENSSKQTL

KTPGKSTFKRSKLILRNMKKSIKNSCPPNPKPAESVQSNSSSFLNFGFANRFRPKG

PRNPPTTWNI"

FEATURES

source

CDS

BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

alignment_scores:
Quality: 3569.00 Length: 685
Ratio: 5.210 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.854

alignment_block:
US-09-600-358A-4 x AF001846 ..

Align seg 1/1 to: AF001846 from: 1 to: 3058

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
42 ATGACCAAGAGAGAAATTCGCAGAAGTTCCTGGATGAGGCCCAAGCAA 91
|||||
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
|||||
92 GAAAATTACTAAAGAGAGGTTTGCCAATGAATTTCTGAAGCTGAAAGGC 141
|||||
34 lnSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
142 AATCTACCAAGTACAGGCAGACAAAACCTATCTACAACTGTGGCTGAG 191
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
|||||
192 AATGCCAAGATATCAAGAAAAACAGATATAGGATATTTGCCCTATGA 241
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
242 TTATAGCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATCCAGCT 291
|||||
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
292 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCCAAGGCTTATAT 341
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetII 117
|||||
342 GCCACCCAGGTCCTTTATCTACAAACCTCCTGGACTTCTGGAGGATGAT 391
|||||
117 eTrrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
392 TTGGGAATATAGTCTCCTTATCATTTGTTATGGCATGCGATGAGTATGAA 441
|||||
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
442 TGGGAAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 491
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
492 CTGGAATTTGGCCCTTCTCTGTATCTCTGTAAGCTGNAAGAAAGAAATC 541
|||||
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
542 TGATTATATAATCAGGACTCTAAAAAGTTTAAGTTCAAATAGTCAAACTCGA 591
|||||
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
592 CTATCTACCAAGTTTCATTACAAAGAAATTGGCCAGACCATGATGTACCTTCA 641
|||||
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
642 TCTATAGACCCATATCTTGACCTCATCTGGGATGTACGTTGTTACCAAGA 691
|||||
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
692 GGATGACAGTGTCCCATATGCCATATCACTGAGTGTCTGGCTGTGGAAGA 741
|||||
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLysAsp 250
|||||
742 CTGGTGTATTATTGTGCTATTGTTGATTATACATGATGTTGCTAAAGAT 791
|||||

251 GlyIlelleProGluAsnPhSerValPheSerLeuIleArgGluMetAr 267
|||||
792 GGGATAATTCCTGAGAACTTCAGTGTTCAGTTTGATCCGGAAATCGG 841

267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
842 GACACAGAGGCTTCATTAGTTCAAAGCAGGAAACAATATGAACGTGCT 891

284 yAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
892 ACAATGCTGATTAAGAACTATTAAGAGACAGATGGATGTTATCAGAGAT 941

301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
942 AAACATTCGGAACAGAGAGTCAGCAAGACATGTTATCCTCGAGAAAA 991

317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
992 TCACACTCTCCAGCAGACTCTTATCTCTTAATTTACCAAAAAATGACCA 1041

334 hLysAlaAlaLysMetMetAsnGlnArgThrLysMetGluIleLys 350
|||||
1042 CAAAGCAGCAAAATGATGAACCAACAAAGGACAAATAATGGAAATCAAA 1091

351 GluSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG1 367
|||||
1092 GAATCTCTTCCTTTGACTTTAGGACTTCGAAATAAGTCGCAAAAGAGA 1141

367 uLeuValLeuHisProLysSerSerThrSerPheAspPheLeuGluL 384
|||||
1142 GCTAGTTTTCACCCCTGCTAAATCAAGCACTTCCTTTGACTTCTGGAGC 1191

384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1192 TAAATTACGTTTTCACAAAAATGCTGACACCAACCATGAATGGCAGACA 1241

401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1242 AGGCATTTCCATAGTTGGGAGCCCTTCAGAAAGCATCAAGTTTGGGA 1291

417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1292 TTTGGGCTCTCTTTGTTTGGGGATGTTCTTAATTCATAACCTGTAATG 1341

434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1342 CAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACACGACCAAAATCA 1391

451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLys 467
|||||
1392 ACTCCTTTGNAATGATACAGCAGAGAGAAACCAAGGAGGTGGACACAA 1441

467 sGluAsnPhSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1442 GGAACACTTTCTTATTTGGAATCTCAACCAACATGATTCCTGTTGTAG 1491

484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1492 AGATGCAGGCTCAAAAAGTAAATGATGCTTTCTTCAGCAGAACTGAATAT 1541

501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLy 517
|||||
1542 TCACTGCCATATGACTCTAACACCAAAATAGTAATGCCCTCTAATGTAAA 1591

517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1592 GCACCATGACTTAGTGCTCTTGGTGATATCTTACATACCTTTAGTGG 1641

534 luAsnProTyrPheSerSerTrpProSerGlyThrSerLysMet 550
|||||
1642 AAAATCCTTATTTTCATCATGCGCTCCAAAGTGGTACCACTTCTAAGATG 1691

551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1692 TCTCTTGATTTACCTGAGAACAGATGGAAGTCTTTTCTCTTCTCTCT 1741

567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1742 GTTCCCAACATCCTCTACATCCCTCTCTCTTATTACAAATTCACATGAT 1791

584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1792 CTTTATCACTGAATTCCTCAACCAATATTTCTCTCTTATGAACGAGAG 1841

601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
|||||
1842 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGAATCCCTCCCTCC 1891

617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
|||||
1892 ACTTCTGTACGGACACTGAATCATTTATTGTGTGGTGGAGAACTGGAG 1941

634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1942 AATTCTCACCAATGTTCCTCAATCTCTATCTCAGCTGTGAAGTAAAA 1991

651 IleGlyThrSerLeuGluTrpGlyThrSerGluProLysLysPheAs 667
|||||
1992 ATTGGAACATCACTGGAATGGGTGGACATCTGAACCAAGAAATTTGA 2041

667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2042 TGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAACTCCGAAAGTCTTA 2091

584 ysSer 585
|||||
2092 AATCA 2096

seq_name: gb_pat:ARI49916
seq_documentation_block:
LOCUS ARI49916 3580 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6228641.
ACCESSION ARI49916
VERSION ARI49916.1 GI:15114507
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3580)
AUTHORS Jallal,B. and Plovman,G.D.
TITLE Diagnosis and treatment of pTP04 related disorders
JOURNAL Patent: US 6228641-A 1 08-MAY-2001;
FEATURES Location/Qualifiers
source
1..3580
BASE COUNT 1213 a 666 c 618 g 1083 t
ORIGIN

alignment_scores:
Quality: 3543.50 Length: 685
Ratio: 5.188 Gaps: 1
Percent Similarity: 99.708 Percent Identity: 99.416

alignment_block:
US-09-600-358A-4 x ARI49916 ..
Align seg 1/1 to: ARI49916 from: 1 to: 3580

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
53 ATGGACCAAGAGAAATTCGAGAAGTTCTGTGATGAGGCCCAAGCAA 102

17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34

|||||
103 GAAATTAATAAGAGGAGCTTTGCCAATGAATTTCTGAAGCTGAAAGGC 152
34 lnsrThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
153 AATCTACCAAGTACAGGAGCAGAAACCTATCCTCAACACTGTGGCTGAG 202
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
203 AAGCCCAACAATATCAAGAAACACAGATATAGGATATTTGCCCTATGA 252
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
253 TTATAGCCGGTAGAACTATCCCTGATACCTCTGTATGAGGATTCAGCT 302
84 yrlleAsnAlaAsnPheIleLysGlyValTyrClyProLysAlaTyrIle 100
303 ACATCAATGCCAACTTCATTAAGGGAGTTATGGACCCAAAGGCTTATATT 352
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetIi 117
353 GCCACCCAGGGTCTTTATCTACAAACCTCTGGACTTCTGGAGGATGAT 402
117 eTrrpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
403 TTGGGAATATAGTGTCTTATCATCTTATGTCATGTCATGGATGATAA 452
134 eTcGlyLysLysCysGluArgTyrTrrpAlaGluProGlyGluMetGln 150
453 TGGGAAGAAAGTGTGAGCGCTACTGGCGTGAGCCAGAGAGATGCAG 502
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
503 CTGGAATTTGGCCCTTCTCTGTATCTCTGTGAGCTGAAAGAGGAATC 552
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
553 TGATTATATAATCAGGACTCTAAAGTTAAGTTCAATAGTGAACATCGAA 602
184 hrIleTyrGlnPheHisTyrLysAsnTrrpProAspHisAspValProSer 200
603 CTATCTACCAAGTTTCATTACAAGAAATGGCCAGACCATGATGATACCTTCA 652
201 SerIleAspProIleLeuGluLeuIleTrrpAspValArgCystyrGlnGl 217
653 TCTATAGACCTATTCTTGAGCTCATCTGGGATGTACGTTGTACCAGA 702
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
703 GGATGACAGTGTCCCATATGCATTCACCTGCAGTGTGGCTGTGAAGGA 752
234 hrGlyValIleCysAlaIleValAspTyrThrTrrpMetLeuLysAsp 250
753 CTGGTGTATTGTGCTATT...GATTATACATGGATGTGTGCTTAAAGAT 799
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
800 GGATTAATCTCGAAGCTTCAGTGTTCAGTTTGTATCGCGGAATGCG 849
267 gThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValT 284
850 GACACAGAGGCTTCATTAGTTCAACCGCAGGAACAATATGAACCTGCTCT 899
284 yAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
900 ACAATGCTGTATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGAT 949
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
950 AAACATCTGGAACAGAGAGTCAAGCAAAAGCATGTGATTCTCAGAAAAA 999
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||

1000 TCACACTCTCCAAGCAGACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1049
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
1050 CAAAACGACAAAAATGATGAACCAACAAGGACAAAAATGGAATCAAA 1099
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
1100 GAATCTTCTCTCTTTGACTTTAGGACTTCTGAAATAAGTGCAAAAGA 1149
367 uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL 384
1150 GCTAGTTTTCACCCCTGCTAAATCAAGCACTTCTTTTGACTTTCTGGAGC 1199
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrrpGlnThr 400
1200 TAAATTTACAGTTTGTACAAAAATGCTGCACACCAACATGAATGGCAGACA 1249
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
1250 AAGGCATTTCCAATAGTTGGGGAGCCTCTTCAGAAGCATCAAAAGTTTGA 1299
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
1300 TTTGGGCTCTCTTTTGTGGGGATGTTCTTAATTTAAACCTGTAAATF 1349
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
1350 CAGCAGAGAGATATTTTAATTCAAAGGTGCCAATAACACGGACCAATCA 1399
451 ThrProPheGluLeuIleGlnGlnArgGluThrLysGluValAspSerLy 467
1400 ACTCTCTTTGAATTTGATACAGCAGAGAGAAACCAAGGAGGTGCACAGCA 1449
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
1450 GGAACCTTTTCTTATTTGGAATCTCAACACATGATTTCTGTTTGTAG 1499
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
1500 AGATGAGGCTCAAAAAGTAATGCATGTTCTTCACGAGAACTGAATTAT 1549
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
1550 TCACCTGCCATATGACTCTAAACACCAAAATACGTAATGCCCTCTAATGTAAA 1599
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProIleuValG 534
1600 GCACCATGACTCTAGTCTCTTGGTGTATATTCTTACATACCTTTAGTGG 1649
534 luAsnProTyrPheSerSerTrrpProProSerGlyThrSerSerLysMet 550
1650 AAAATCCTTATTTTTCATCATGCTCCCAAGTGGTACCAGTTCTAAGATG 1699
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerle 567
1700 TCTCTTTGATTTTACCTGAGAAGCAAGATGGAACGTGTTTTTCTCTCTCT 1749
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
1750 GTTGCCAAACATCTCTACATCCCTCTTCTTATTTACAATTCACATGATT 1799
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
1800 CTTTATACATGAATTTCCAAACCAATATTTTCTCCTACTATTGAACCCAGGAG 1849
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProProPr 617
1850 TCAGCTGTACTAGCAACTGCTCCAAGGATAGATGATGAAATCCCCCTCC 1899
617 oLeuProValArgThrProGluSerPheIleValValGluGluAlaGlyG 634
1900 ACTTCCTGTACGGACACCTGAATCATTTATTTGTGGTTGAGGAAGCTGGAG 1949

634 luPheSerProAsnValProLysSerLeuSerSerAlaValLysValLys 650
|||||
1950 AATTCTCACCAAAATGTTCCCAAACTTATCTCAGCTGTGAAGTAAAA 1999
651 lIleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
2000 ATTGGAACATCAGTGGGATGGGTGGAACATCTCTGAACCAAGAAATTTGA 2049
667 pASpSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2050 TGACTCTGTGATACTTACCAACAAGAGGTGTAACCTCCGAAGTCTCA 2099
684 ySser 685
|||||
2100 AATCA 2104

seq_name: gb_pr:AF150732

seq_documentation_block:

LOCUS AF150732 2877 bp mRNA linear PRI 04-MAY-2000
DEFINITION Homo sapiens protein tyrosine phosphatase (PTP) mRNA, complete cds.

ACCESSION AF150732

VERSION AF150732.1 GI:7688662

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2877)

Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,

Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.

A novel gene expressed in human adrenal gland

Unpublished

2 (bases 1 to 2877)

Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,

Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.

Direct Submission

Submitted (12-MAY-1999) Chinese National Human Genome Center at

Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,

Shanghai 201203, P. R. China

FEATURES

source

1. .2877

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue.type="adrenal gland"

1. .2877

/gene="ptp"

71. .2470

/gene="ptp"

/codon.start=1

/product="protein tyrosine phosphatase"

/protein_id="AAF67472.1"

/db_xref="GI:7688662"

/translation="MDQREILQKFLDEAQSKKITKEEFANEFKLKROSTKYKADKTY

PTVAEKPNKIKNRYKDIPLVDYSKVELSLITSDSSYINANFKYVGEAVIAT

QGPLSTLLDFRMWLEYSLIIVMACMEYEMKKCRPYWAEPGEMQLFEGFVSVC

EAERKSDYIIRLTKVKNSERTYIQHYKNPDPHVPSSIDPILELWVRCQED

DSVPICISACGKRGTVICADYTWMLLKDGLIIPENFSVSLIREMTRQPSLVOTQ

EYELVYNAVLEKFRQMDVIRDKHSGTESQAKHCIPKKNHTLQADSYSPNLPKSTTK

AAKMNQKRTMEIKRESSDFRTSEISAKEDVSLAPENQAFSCISEVNSFAPKMLPH

HEMPAKAFQIVGSSEASKLNWASLCGMNSNLVCSRKIPNRCQYPDQINSFELIQ

QENQGDGSKTFLFESQPHDSCFVEMQAKVMHVSSAELNYSLPYDSKHQIINASNVK

HDSSALGVYSYIPIVENPYSSWPSTSSKMSLDLPKQDGVFPSSLPTSTSL

FSYINSHDSLANSVNTINSSRIESESVALAPRIDDEIPLHFLYGTPESTIVVEAG

SSPPPLPERTLESFLADECMQAQSIETYSYPTDMENSTSSKQILKTPGKSFTR

SKSLKILRMKKSICNSCPPNKPAPESVQSNSSSFLNFGFANRESKPGPRNPPTWN

I"

BASE COUNT

ORIGIN

944 a 583 c 523 g 827 t

alignment_scores:

Quality: 2977.50 Length: 687

Ratio: 4.660 Gaps: 11

Percent Similarity: 93.013 Percent Identity: 89.229

alignment_block:

US-09-600-358A-4 x AF150732 ..

Align seg 1/1 to: AF150732 from: 1 to: 2877

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||

71 ATGGACCAAGAGAGAAATCTCAGAAATTCCTGGATGAGGCCCAAGCAA 120

|||||

17 sLYSTleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34

|||||

121 GAAATTAATAAGAGAGGATTTGCCAATGAATTTCTGAAGCTGAAAGGC 170

|||||

34 lNserThrLysTyLysAlaAspLysThrTyProThrThrValAlaGlu 50

|||||

171 AATCTACCAAGTACAGGCACACAAACCTATCTTACAACTGTGGCTGAG 220

|||||

51 AsnAlaLysAsnIleLysLysAsnArgTyLysAspIleLeuProTyAs 67

::: |||||

221 AAGCCCAAGAATATCAAGAAAAACAGATATAAGGATATTTGCCCTATCA 270

|||||

67 pTySerArgValGluLeuSerLeuIleThrSerAspGluAspSerSer 84

|||||

271 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATTCGAGCT 320

|||||

84 yRIleAsnAlaAsnPheIleLysGlyValTyGlyProLysAlaTyIle 100

|||||

321 ACATCAATGCCAACTTCATTAAAGGAGTTTATGGACCCGAGGCTTATAT 370

|||||

101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117

|||||

371 GCCACCCAGGGTCCTTTATCTACAACCTCTCGGACTTCTGGAGGATGAT 420

|||||

117 eTrpGluTySerValLeuIleIleValMetAlaCysMetGluTyGluM 134

|||||

421 TTGGGAATATAGTGTCCTTATCATTTGTTATGGCATGGAGTATGAAA 470

|||||

134 etGlyLysLysCysGluArgTyTrpAlaGluProGlyGluMetGln 150

|||||

471 TGGGAAGAAAGTGTGAGCGCTACTGGGCTGAGCCAGAGAGATGCAG 520

|||||

151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLys 167

|||||

521 CTGGAATTTGGCCCTTCTCTGTATCTCTGTAAGCTGAAAAAGGAAATC 570

|||||

167 rASpTyIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184

|||||

571 TGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAATCTGAA 620

|||||

184 hrIleTyGlnPheHisTyLysAsnTrpProAspHisAspValProSer 200

|||||

621 CTATCTACCACTTCATTACAGAATTTGGCCAGACCATGATGTACCTTCA 670

|||||

201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyTrpGln 217

|||||

671 TCTATAGACCCCTATTCTTGGAGCTCATCTGGGATGTCCTGTTACCCAAGA 720

|||||

217 uASpAspSerValProIleCysIleIleHisCysSerAlaGlyCysGlyArgT 234

|||||

721 GGATGACAGTGTCCCATATCATCTCACTGAGTGTGGCTGTGGGAAGGA 770

|||||

234 hrGlyValIleCysAlaIleValAspTyTrpMetLeuLeuLysAsp 250

|||||

771 CTGGTGTATTGTGCTATT...GATTATACATGGATGTTGCTAAAAAGAT 817

|||||

251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267

|||||

818 GGGATAATTCCTGAGAACTTCAGTGTGTTTTCAGTTTGTATCCGGGAAATCGC 867

267	gThrGlnArgProSerLeuValGlnThrGlnGluGlnThrGluLeuVal	284
868	GACACAGAGGCTTCATTAGTTCAAAGCGCAGGACAAATATGAATCGTCT	917
284	yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleAqAsp	300
918	ACAATGCTGTATTAGAATATTTAAGAGACAGATGGAATGTTATCAGAGAT	967
301	LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs	317
968	AAACATTCTGGAACAGAGAGTCAAGCAAAGCATTTGTTCTCAGAAAAA	1017
317	nhIsthrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT	334
1018	TCACACTCTTCAAGCAGACTCTTATCTCTCTAAATTTACCAAAAGTACCA	1067
334	hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys	350
1068	CAAAAGCAGCAAAATGATGAACCAAGAGCAAAATGGAATCCAA	1117
351	GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluG	367
1118	GAATCTTCTCTTGACTTTTAGGACTCTGAATATAGTGCAAAAGAA	1167
367	uLeuValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluL	384
1168	CGTAAGTTTGGCCCCG...GAAACACAGGCTTTTTTTTGTGCTTTCGGAAG	1214
384	euaSntSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr	400
1215	TAAATTCAGTTTTCAAAATGTTG...CCCCACCATGAATGCCAGCA	1261
401	LysAlaPheProIleValGlyGluPro...LeuGlnLysHisGlnSerLeuA	417
1262	AAGGCTTTCCAAATAGTGGGA...AGCTCTTCAGACGATCAAGT...T	1305
417	sPLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsn	433
1306	GAATTTGGGCTCTCTTTCTTTAGGAATGTCFA...ATTCTAACTCTGTAT	1352
434	AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe	450
1353	GCAGCAGAAAGATATTAATCCAGGTGCCAAT...ACCGGACCAATC	1399
450	rThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerL	467
1400	AACTCT.TTTGAATTGATCCAGCAA...GAGAACCAAGAGGTGACAGCA	1445
467	ysGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheVal	483
1446	AGAAACATTTCTTAFTT...GAATCTCAACACATGATCTCTGTTTGTGA	1492
484	GluMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTy	500
1493	GAGATGCAGGCTCAAAAAGTAATGCAATGTTCTTCAGCAGNACTGAATTA	1542
500	rSerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal	517
1543	TTCACTGGCATATGACTCTAAACACCAAAATACGTAATGCCTCTAATGTA	1592
517	ysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVal	533
1593	AGCACCATGACTCTAGTCTCTGGTGATATCTTACATACCTTTAGTG	1642
534	GluAsnProTyrPheSerSerTriProProSerGlyThrSerSerLysme	550
1643	GAANAATCTTATTTTTCATCATGGCCCTCAAGTGGTACCAGTCTTAGAT	1692
550	tSerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerL	567
1693	GTCTCTTGATTTACCTGAGAAGCAAGATGGAACGTGTTTTCTCTTCTCTC	1742

567 euLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSer 583
 1743 TGTTGCCAACATCCCTCTACATCCCTCTCTCTATTATCAATTCACATGAT 1792

584 SerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnG1 600
 1793 TCTTTATCACTCAATCTCCACCAATATTTCTCTACGATTATGACAGGA 1842

600 uSerIaValLeuAlaThrAlaProArgIleAspAspGluIleProProp 617
 1843 GTCAGCTGTACTAGCAACCTGCTCCAAGGATAGATGATGAAT_CCCCTC 1891

617 toLeuProValarg_ThrProGluSerPheIleValValGluGluAlaG1 633
 1892 CACTTCTCTGACGGGACCTGAATCATTTATTTGTGTTGAGGAAGCTGG 1941

633 yGluPheSerProAsnValProLysSerLeuSerSerAlaValLysValL 650
 1942 AGAATTTCTCACCAAAATGTTCCCAATCTTATCTCAGCTGTGAAGTAA 1991

650 ysIleGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPhe 666
 1992 AAATTGGACATCACTGGAATGGGTGGGAACATCTGAACCAAGAAATTT 2041

667 AspAspSerValIleLeuArgProSerLysSerValLysLeuArgSerPr 683
 2042 GATGACTCTGTGATCACTTATAGACCACAGAGGTGTAATCTCGAAGTCC 2091

683 oLysSer 685
 2092 TAAATCA 2098

seq_name: gb_ro:MUSPROTYPH

seq_documentation_block:
 LOCUS MUSPROTYPH 2734 bp mRNA linear
 DEFINITION Mouse protein tyrosine phosphatase (70zpep) mRNA, com
 ACCESSION M90388
 VERSION M90388.1 GI:200522
 KEYWORDS protein-tyrosine phosphatase.
 SOURCE Mus musculus cDNA to mRNA.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
 1 (bases 1 to 2734)
 AUTHORS Matthews,R.J., Bowne,D.B., Flores,E. and Thomas,M.L.
 TITLE Characterization of hematopoietic intracellular prote
 and another enriched in proline-, glutamic acid-, ser
 threonine-rich sequences
 Mol. Cell. Biol. 12, 2396-2405 (1992)
 JOURNAL 92236615
 MEDLINE Location/Qualifiers
 FEATURES
 source 1..2734
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 gene 1..2734
 /gene="70zpep"
 CDS 113..2521
 /gene="70zpep"
 /EC_number="2.7.1.112"
 /codon_start=1
 /product="protein tyrosine phosphatase"
 /protein_id="AAA39994.1"
 /db_xref="GI:200523"
 /translation="MDQREILQQLLKEAQKKLNSEEFASEFL
 PTTVAQRPNKTKNRYKIDILPDHSLVELSLTSDSSYINAA
 QGPLSTLLDFWRIMETRIIYQFHKWPDHVPDSSINAE
 EAEEKKSDYKTRLKAKENETRIIYQFHKWPDHVPDSSINAE
 DQVETLCHSAGCGRTGICAVDYTWMLKDGIIIPKFSVFNLI
 EYELVYSAVLELFRKHMVDISDNLHGREIQACSIPESQLTW
 DKVKTNOHSGKGAESTGGSSIGLRTSTMNAEELVLHSAKSL
 KAVTTRNGKASASDVCPQKQVOSIDFCSPTLFGCPSALPTN

KSTPFELIOQRKINDIADVAGDFSCLESQLEHYSLRELQVORVAHVSSBELNYSLPGA
CDASCVPRLSPGALRVHLVYLSLAEDPYFSSPPNSADSKMSFDLPEKODGATSPGALL
PASSTSFYVSNPHDSLWNTLTFSPLNOETAVEAPSRDDDEIPPLPRTPESE
IVVEEAGEPSRVVTESLPLVTFEGASPEGSTSEKSHDSVGFPSKNVYKLSRPSKDR
HODGSPPLPRTLESFLAEDCQAOAVQTSSTYPETENSTSSKQTLRTPKRS
FRKSLKIFRNKMSVCSNSSPSKPRVQPRKNSSSLNFGFGRFSPKPGPRNPPS
ANW*

BASE COUNT 787 a 629 c 601 g 717 t
ORIGIN

alignment_scores:

Quality: 2356.50 Length: 692
Ratio: 4.028 Gaps: 6
Percent Similarity: 84.538 Percent Identity: 68.497

alignment_block:

US-09-600-358A-4 x MUSPROTYPH ..

Align seg 1/1 to: MUSPROTYPH from: 1 to: 2734

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
113 ATGGACCAAGAGAAATCTCGACCACTACTGAAAGAACGCCAGAAAA 162
17 sLysileThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
163 GAAACTCAACAGTCAGGAGTTTGCAGTGAATTTCTGAAGCTGAAAGGC 212
34 lnSerThrLysTyLysAlaAspLysThrTyProThrThrValAlaGlu 50
|||||
213 AATCAGCAAGTACAGGCCGACAAAAATCTATCTACAACTGTGGCTCAG 262
51 AsnAlaLysAsnIleLysLysAsnArgTyLysAspIleLeuProTyAs 67
::: |||||
263 AGGCCCAAGAAATATCAAGAAAAACAGATACAAAGGATATTTGCCCTATGA 312
67 pTySerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||: |||||
313 TCACAGCTGTGACAGTGTCTGTGTAACCTCCGATGAGGATCCAGTT 362
84 yrIleAsnAlaAsnPheIleLysGlyValTyGlyProLysAlaTyIle 100
|||: |||||
363 ATATCAATGCCAGCTTTTAAAGGCTGTCTATGACCAAGGCTTATATT 412
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117
|||||
413 GCTACCCAGGGTCTTTATCTACAACTCTCCTGGACTTCTGGAGGATGAT 462
117 eTrpGluTySerValLeuIleIleValMetAlaCysMetGluTyGlu 134
|||||
463 TTGGAGTACCGCATCTTGGTCATTTGTCATGGCATGTATGGAGTTGAAA 512
134 etGlyLysLysCysGluArgTyTrpAlaGluProGlyGluMetGln 150
|||||
513 TGGGAAAAAATAATGTAGCGTTTATTGGGCCGCAACAGGAGAAACGCG 562
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||: |||||
563 CTGCAATTTGGCCCTTTTCTATATCTGTGAAGCTGAGAAAAAGAAATC 612
167 rAspTyIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
613 TGATATATAATCAGGACTCTGAAGGCCAAGTTCAATAATGAAGAACTGAA 662
184 hrIleTyGlnPheHisTyLysAsnTrpProAspHisAspValProSer 200
|||||
663 TTATTTACCACTTTTCAATATAAAGATTGGCCAGACCATGATGTGCCITCA 712
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyGlnGl 217
|||||
713 TCTATAGACCCCTATTCTCAGCTCATCTGGGATATGGCTTGTACCAAGA 762

217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
763 AGATGACTGTCTTCTATATCATCTCACTGAGTCCGCGCTCGGGAAGA 812
234 hrGlyValIleCysAlaIleValAspTyThrTrpMetLeuLeuLysAsp 250
|||||
813 CAGGTGTCTATTGTGCT..GTTGATTATACATGATGCTGCTGAAAGAT 859
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
860 GGGATAATCTCTAAGAACTTCACTGTTTAAATTTGATTACAGGATGCG 909
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyGluLeuValT 284
|||||
910 AACACAGAGGCTTCGCTAGTTTCAAACTCAGGAACAGTACGAACTGGTCT 959
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgasp 300
|||||
960 ACAGTGTCTGTATTAGAGCTGTTTAAAGAGGCACATGGATGTTATCTCTGAT 1009
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
::: |||||
1010 AATCACCTTGAAGAGAGATTCAGACACATGCTCAATTCCTGAAACAAG 1059
317 nHisThrLeuGlnAlaAspSerTySerProAsnLeuProLysSerThr 334
: |||||
1060 CCTCAGCGTAGAAGCTGACTCTTGTCTCTGGAATTTACAAAAACGCCA 1109
334 hrLysAlaAlaLysMetMetAsnGln.....GlnArgThrLysMet 347
::: |||||
1110 TGAGGGATGTGAAAAACGACAAACAGCATAGCAACAAAGGGGCTGAAGCG 1159
348 GluIleLysGluSerSerSerPheAspPheArgThrSerGluIleSerAl 364
|||||
1160 GAGAGCACTGGAGGCTCTTCCCTTGGCCTTAGGACTTCTACGATGAATGC 1209
364 aLysGluGluLeuValLeuHisProAlaLysSerSerThrSerPheAsp 381
|||: |||||
1210 CGAGGAAGAGTTGGTTTGCACCTCGCTTAATCAAGCCCTCTTTTAAC 1259
381 heLeuGluLeuAsnTySerPheAspLysAsnAlaAspThrThrMetLys 397
|||||
1260 GTTTAGAGCTAAACTGGCGGTGTAAACAAAGGCTGTATACCAAGGAAAC 1309
398 TrpGlnThrLysAlaPheProIleValGlyGluProLeuGlnLysHisGl 414
|||||
1310 GGGCAGGCAAGGCTTCTCCAGTCTGGGAGAGCCCTTCAGAGATATCA 1359
414 nSerLeuAspLeuGlySerLeuPheGluGlyCysSerAsnSerLysp 431
|||||
1360 AAGCTGATTTTGGTTCCATGTTTGGGCTCGTGTCTAGTGTCTGCTGCG 1409
431 roValAsnAlaAlaGlyArgTyPheAsnSerLysValProIleThrArg 447
|||||
1410 CCATAAACACACAGCGACAGGTATCACATTCATAAGGGCGCGGTAAACGG 1459
448 ThrLysSerThrProPheGluLeuIleGlnArgGluThrLysGluVa 464
|||||
1460 ACCAAATCACTCCCTTTGAATGATTCAGCAGAGAAAAACAAATGACTT 1509
464 lAspSerLysGluAsnPheSerTyLeuLeuSerGlnProHisAspSerc 481
: |||||
1510 GGCCGTGGGAGACGGTTTTCATGCTGGAATCTCAGCTGCATGAGCAT 1559
481 ysPheVal...GluMetGlnAlaGlnLysValMetHisValSerSerAla 496
::: |||||
1560 ACAGTCTCAGGAGCTGCAGGTGCAAGAGTGGCCCATGTTTCTTCAGAA 1609
497 GluLeuAsnTySerLeuProTyAspSerLysHisGlnIleArgAsnAl 513
|||||
1610 GAGCTGAATTTATCTACCTGCCT.....GGTGCCTGTGTGATGC 1644
513 aserAsnValLysHisHisAspSerSerAlaLeuGlyValTySerTyI 530

```
||||| ||| ::::: ::::: ||||| ||::: |||
1645 GTGGTGTGCCCCGACAGCCCGCGCTTTAGAGTGCATCTGTACA 1694
530 leProLeuValGluAsnProTyrPheSerSerTrpProSerGlyThr 546
||||| ::::: ::::: ||||| ||||| :::::
1695 CATCTTACGGGAAGATCTTATTTTCATCATCCCTCCGATATGCT 1744
547 SerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrValPh 563
||||| ::::: ::::: ||||| ||||| :::::
1745 GATTCGAAGATGCTTTTGTATGCTGCTGAGAACAGAGATGGACCACTTC 1794
563 eProSerSerLeuLeuProThrSerThrSerLeuPheSerTyrTyra 580
||||| ::::: ::::: ||||| ||||| :::::
1795 CCTGGCGCTATTGCGGCTCTTCTTACAACTCTCTTTTATATAGCA 1844
580 snSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeu 596
||| ::::: ||||| ::::: ||||| :::::
1845 ACCACGACACTCCCTAGTGATGAACACTCTGACCACTTTTCCACCG 1894
597 LeuAsnGlnGluSerAlaValLeuAlaThrAlaProArgIleAspAspG 613
||||| ::::: ||||| ||| ::::: |||||
1895 TTAACACAGACAGACCTGTAGAGCTCTTCTCGGAGGACAGATGATGA 1944
613 uileProProLeuProValArgThrProGluSerPheIleValValG 630
||||| ::::: ||||| ||||| ::::: |||||
1945 AATCCCGCCGCCACTCCCTGAACGACACCCGAGTCTTTTATTGTGTTG 1994
630 lucluAlaGlyGluPheSerProAsnValProLysSerLeuSerAla 646
||||| ::::: ||||| ||||| ::::: |||||
1995 AGGAAGCCGGAGAGCTCTCACCAGTGTACCGAATCCCTTA.....CCT 2038
647 ValLysValLysIleGlyThrSerLeuGluTrpGlyThrSerGluPr 663
::: ||| ::::: ||||| ||| ::::: |||||
2039 CTGGTGTAACTTTGGAGCATCACCAGATGCAGTGGGACATCTGAAT 2088
663 oLysLysPheAspAspSerValIleLeuArgProSerLysSerValLysL 680
||||| ::::: ||||| ::::: ||||| ::::: |||||
2089 GAAGAGC...CATGACTCTGTAGGTTTACACCAAGCAAGATGTGAAC 2135
680 euArgSerProLysSerGlyLysAsn 688
||||| ::::: ||||| ::::: |||||
2136 TCCGAAGTCCCAATCAGATCGACAT 2161
```

seq_name: gb_pr:AL137856

seq_documentation_block: 143618 bp DNA linear PRI 17-JUL-2001
LOCUS AL137856 Human DNA sequence from clone RP5-107303 on chromosome 1p13.1-13.3,
DEFINITION complete sequence.

ACCESSION AL137856

VERSION AL137856.24 GI:14970654

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 143618)

Glithero,R.

Direct Submission

Submitted (17-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Jul 19, 2001 this sequence version replaced gi:14586004.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, ENBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>

RP5-107303 is from the library RPCI-5 constructed by the group of

Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pcYPAC2

This sequence is the entire insert of clone RP5-107303 The true

right end of clone RP11-215C20 is at 30467 in this sequence. The

true right end of clone RP11-324J2 is at 119169 in this sequence.

FEATURES

Source

1..143618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p13.1-13.3"
/clone="RP5-107303"
/clone_lib="RPCI-5"
257..322
/note="MERSA repeat: matches 41..109 of consensus"
763..1053
/note="AluX repeat: matches 1..312 of consensus"
1348..1698
/note="L2 repeat: matches 2070..2417 of consensus"
1919..1969
/note="L2 repeat: matches 2655..2705 of consensus"
2305..2446
/note="MIR repeat: matches 40..176 of consensus"
5652..5927
/note="L1ME repeat: matches 5524..5813 of consensus"
6295..6428
/note="BC200 repeat: matches 2..147 of consensus"
6653..6700
/note="8 copies 6 mer acacac 97% conserved"
7925..7980
/note="28 copies 2 mer aa 75% conserved"
8071..8369
/note="AluX repeat: matches 1..296 of consensus"
9048..9342
/note="AluX repeat: matches 1..298 of consensus"
9793..10559
/note="L1MC1 repeat: matches 5526..6320 of consensus"
10574..10730
/note="MIR repeat: matches 95..262 of consensus"
10759..10907
/note="MIR repeat: matches 2..139 of consensus"
10908..11219
/note="AluJo repeat: matches 1..296 of consensus"
11220..11315
/note="MIR repeat: matches 139..237 of consensus"
12210..12464
/note="AluSg repeat: matches 1..291 of consensus"
12916..13149
/note="MERS8B repeat: matches 70..341 of consensus"
13150..13435
/note="AluX repeat: matches 1..298 of consensus"
13458..13805
/note="L1MD2 repeat: matches 4851..5213 of consensus"
13806..14115
/note="AluJb repeat: matches 1..299 of consensus"
14118..14430
/note="AluX repeat: matches 1..312 of consensus"
14431..14563
/note="L1MD2 repeat: matches 5213..5314 of consensus"


```
repeat_region 14565..14859
/notes="AluSq repeat: matches 1. .294 of consensus"
repeat_region 14861..14888
/notes="14 copies 2 mer ta 92% conserved"
repeat_region 14910..15215
/notes="L1MD2 repeat: matches 5302. .5609 of consensus"
repeat_region 15230..15560
/notes="MER2 repeat: matches 2. .345 of consensus"
repeat_region 15561..16146
/notes="L1MD2 repeat: matches 5643. .6210 of consensus"
repeat_region 16270..16479
/notes="MIR repeat: matches 48. .262 of consensus"
repeat_region 16840..17032
/notes="MER3 repeat: matches 2. .207 of consensus"
repeat_region 17472..17501
/notes="15 copies 2 mer ac 100% conserved"
misc_feature 1839..19765
/notes="CpG Island"
/evidence=not_experimental
20385..20649
/notes="AluJo repeat: matches 1. .297 of consensus"
repeat_region 20642..20734
/notes="FLAM_A repeat: matches 4. .116 of consensus"
repeat_region 22760..23065
/notes="AluSx repeat: matches 1. .306 of consensus"
repeat_region 23275..23506
/notes="AluJb repeat: matches 1. .225 of consensus"
repeat_region 23540..23584
/notes="L2 repeat: matches 2647. .2694 of consensus"
repeat_region 23590..23671
/notes="AluJ/FLAM repeat: matches 1. .83 of consensus"
repeat_region 23714..24013
/notes="AluSx repeat: matches 1. .297 of consensus"
repeat_region 24532..24817
/notes="AluJb repeat: matches 17. .293 of consensus"
repeat_region 25033..25166
/notes="FLAM_C repeat: matches 11. .143 of consensus"
repeat_region 25261..25549
/notes="AluSx repeat: matches 1. .289 of consensus"
repeat_region 25675..25877
/notes="MLTJ repeat: matches 2. .214 of consensus"
repeat_region 25998..26083
/notes="Charliel repeat: matches 3. .92 of consensus"
repeat_region 27261..27556
/notes="AluSx repeat: matches 3. .289 of consensus"
repeat_region 28123..28175
/notes="L2 repeat: matches 2678. .2733 of consensus"
repeat_region 28395..28708
/notes="AluSg repeat: matches 1. .310 of consensus"
repeat_region 29346..29766
/notes="L2 repeat: matches 2085. .2448 of consensus"
repeat_region 29767..30052
/notes="AluSg repeat: matches 1. .287 of consensus"
repeat_region 30053..30345
/notes="L2 repeat: matches 2448. .2750 of consensus"
repeat_region 31191..31486
/notes="AluSg repeat: matches 1. .296 of consensus"
repeat_region 31804..31932
/notes="AluJb repeat: matches 1. .127 of consensus"
repeat_region 32054..32075
/notes="11 copies 2 mer aa 100% conserved"
repeat_region 32503..32780
/notes="AluJo repeat: matches 1. .288 of consensus"
repeat_region 32791..32817
/notes="9 copies 3 mer tgt 96% conserved"
repeat_region 32817..33099
/notes="AluSg repeat: matches 1. .282 of consensus"
repeat_region 33156..33450
/notes="AluJb repeat: matches 6. .301 of consensus"
repeat_region 33747..34055
/notes="AluA5 repeat: matches 1. .309 of consensus"
repeat_region 34584..34775
/notes="L1ME repeat: matches 5644. .5823 of consensus"
```

```
repeat_region 34786..34930
/notes="FLAM_C repeat: matches 12. .142 of consensus"
repeat_region 34931..35225
/notes="L1ME repeat: matches 5265. .5563 of consensus"
repeat_region 35257..35727
/notes="L1ME repeat: matches 4762. .5233 of consensus"
repeat_region 35743..36021
/notes="AluJb repeat: matches 1. .287 of consensus"
repeat_region 36340..36488
/notes="L1PA15 repeat: matches 6009. .6157 of consensus"
repeat_region 37329..37776
/notes="L1MC5 repeat: matches 7275. .7680 of consensus"
repeat_region 38067..38121
/notes="L2 repeat: matches 2681. .2735 of consensus"
repeat_region 38346..38564
/notes="L1MB5 repeat: matches 5921. .6142 of consensus"
repeat_region 39036..39268
/notes="MIR repeat: matches 20. .262 of consensus"
repeat_region 39278..39564
/notes="AluSx repeat: matches 1. .294 of consensus"
repeat_region 39575..39657
/notes="AluJo/FRAM repeat: matches 206. .289 of consensus"
repeat_region 41653..41957
/notes="AluJo repeat: matches 5. .293 of consensus"
repeat_region 42324..42634
/notes="AluY repeat: matches 1. .311 of consensus"
repeat_region 43028..43056

alignment_scores:
    Quality: 1490.50      Length: 345
    Ratio: 4.985         Gaps: 2
    Percent Similarity: 86.667    Percent Identity: 86.087

alignment_block:
US-09-600-358A-4 x AL137856/rev ..

Align seg 1/1 to reverse of: AL137856 from: 1 to: 143618

306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
:::|||||
45710 AAGAGTCAAGCAAGCAATGTATTCTCGAGAAAATCACACTCTCCAAGC 45661
322 aASPserTyrsSerProAsnLeuPro.Lys..... 331
|||||
45660 AGACTCTATTCTCTTAATTTACCAAAAGGTGAGTACATAATTTATATAT 45611
331 ..... 331

45610 CTAAATGTTGACTAATTGCTGCTGGAGATGAATCTCATGCTTATACATC 45561
331 ..... 331

45560 ATTGTGATGCTAGGAAGCGGCTGTTGGGTATGGTGTTCATCATTCAGC 45511
332 .....SerThrLysAlaAlaLysMetMetAsnGlnGln 343
|||||
45510 TTTCTTTTGTGTTAGTACCACAAAGCAGCAAAATGATGACCAACAA 45461
344 ArgThrLysMetGluLeuLysGluSerSerSerPheAspPheArgThrSe 360
|||||
45460 AGGACAAAATGGAATCAAAAGAACTCTTCTCTTTGACTTTAGGACTTC 45411
360 xGluLeuSerAlaLysGluGluLeuValLeuHisProAlaLysSerSerT 377
|||||
45410 TGAATTAAGTGCAAAAGAGAGCTAGTTTTCACCCCTGCTAAATCAAGCA 45361
377 hrSerPheAspPheLeuGluLeuAsnTyrsSerPheAspLysAsnAlaAsp 393
|||||
45360 CTTCCTTTTGACTTCTTCGAGCTAAATTTACAGTTTGGACAAAATGCTGAC 45311
394 ThrThrMetLysTrpGlnThrLysAlaPheProIleValGlyGluProLe 410
|||||
45310 ACAACCATGAATGGCAGACAAAGGCATTTCCAATAGTTGGGGAGCCTCT 45261
```

410 uGlnLysHisGlnSerLeuAspLeuGlySerLeuLeuPheGluGlyCys 427
|||||
45260 TCAGAGCATCAAGTTGGATTTGGCTCTCTTTGTTGGAGGATGT 45211
|||||
427 erAsnSerLysProValAsnAlaAlaGlyArgTyrPheAsnSerLysVal 443
|||||
45210 CTAATTCTAAACCTGTAATGCAGCAGGAAGATATTTAATTCAAAGGTG 45161
|||||
444 ProIleThrArgThrLysSerThrPropheGluLeuIleGlnGlnArgGI 460
|||||
45160 CCAATAACACGACCAATCAACTCTTTTGAATTGATACAGCAGAGA 45111
|||||
460 uThrLysGluValAspSerLysGluAsnPheSerTyrLeuGluSerGlnP 477
|||||
45110 AACCAAGGAGTGGACAGCAGGAAGAACTTTCTTATTGGATCTCAAC 45061
|||||
477 roHisAspSerCysPheValGluMetGlnAlaGlnLysValMetHisVal 493
|||||
45060 CACATGATCTCTGTTGTAGAGATGCAGGCTCAAAAAGTAATGCATGTT 45011
|||||
494 SerSerAlaGluLeuAsnTyrSerLeuProTyrAspSerLysHisGlnII 510
|||||
45010 TCTTCAGCAGAACTGAATTATTCACCTGCCATATGACTCTAAACACCAAT 44961
|||||
510 eArgAsnAlaSerAsnValLysHisAspSerSerAlaLeuGlyValT 527
|||||
44960 ACCTAATGCTCTTAAAGTAAAGCACCATGACTCTAGTCTCTGGTGTAT 44911
|||||
527 yrSerTyrIleProLeuValGluAsnProTyrPheSerSerTyrProPro 543
|||||
44910 ATTCTTACATACCTTTAGTGAAATCTCTTATTTTTCATCATGCGCTCCA 44861
|||||
544 SerGlyThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGI 560
|||||
44860 AGTGGTACCAGTCTAAGATGCTCTTGATTACCTGAGAAGCAAGATGG 44811
|||||
560 yThrValPheProSerSerLeuLeuProThrSerSerThrSerLeuPheS 577
|||||
44810 AACTGTTTTCCTCTCTCTCTGTTGCCAACATCCTCTACATCCTCTCT 44761
|||||
577 erTyrTyrAsnSerHisSerSerLeuSerLeuAsnSerProThrAsnIle 593
|||||
44760 CTTATTACAATTCACATGATCTTTATCTACTGAATTCCTCAACCAATATT 44711
|||||
594 SerSerLeuLeuAsnGlnGluSerAlaValLeu 604
|||||
44710 TCTCTACTATTGAACCAAGGAGTCAGCTGTACTA 44678
|||||

seq_name: gb_pr:HUMPTPG1

seq_documentation_block:
LOCUS HUMPTPG1 2938 bp mRNA linear PRI 23-JUN-1999
DEFINITION Human mRNA for protein-tyrosine-phosphatase GI, complete cds.
ACCESSION D13380
VERSION D13380.1 GI:220033
KEYWORDS PEST-domain phosphatase; leukocyte common antigen related molecule;
protein tyrosine phosphatase GI
SOURCE Homo sapiens adult colon cDNA to mRNA, clone_lib:Clontech.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2938)
AUTHORS Takekawa,M., Itoh,F., Hinoda,Y., Arimura,Y., Toyota,M., Sekiya,M.,
Adachi,M., Imai,K. and Yachi,A.
TITLE Cloning and characterization of a human cDNA encoding a novel
putative cytoplasmic protein-tyrosine-phosphatase
JOURNAL Biochem. Biophys. Res. Commun. 189 (2), 1223-1230 (1992)
MEDLINE 93112015
AUTHORS 2 (bases 1 to 2938)
Takekawa,M.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-1992) Mutsuhiro Takekawa, Sapporo Medical

University, Internal Medicine (Section1); South 1, West 16,
Chuo-ku, Sapporo 060, Japan
(E-mail:kimura&apmed.ac.jp@vax2.nlm.nih.gov,
Tel:011-611-2111(ex.3211), Fax:011-613-1141)
Submitted (09-OCT-1992) to DDBJ by:
Mutsuhiro Takekawa

Dept. of Internal Medicine Section I
Sapporo Medical College
South 1, West 17
Sapporo, 060
Japan

Phone: 011-611-2111 x3211
Fax: 011-613-1141.

FEATURES
source

1. 2938 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="colon"
/clone_lib="Clontech"
/dev_stage="adult"
20. 2362
/codon_start=1
/product="protein tyrosine phosphatase GI"
/protein_id="BAA03648.1"
/db_xref="GI:220034"

CDS

/translation="MEQVEILRKFIQRVQAMKSPDHNGEDNFARDFMRLRLSTKYRT
EKIYPTATGEKEENVKNKYKDILPFDHSRVKLTLPDSQSDYINANFIKGVYGPKA
YVATGGPLANTVIDFNMWYENVVIVACREFEMRKCKERYWPLXGEDPITPAPE
KISCDEQARDYFIRTLLEFQNESRLYQFHYVWPDHVPSSFDSILDMISLMRK
YOEHDVPICHCAGCGTGAICAIDYTNLLKAGKIPEEFNVFNLIQEMTKQRHSA
VOTKEQYELVHRAIAQLEKQLQYIEHGAQKIADGVNEINTENNVSIEPKQDSP
PKPRTSCLVEGDAKEILQPEPPHVPPIITPSPSAFTVTVTWQDNDRYHPKPV
LHMVSSEOSADLNRNYSKTELPCKNESTEOIDKLENLSEFKKVPLOEGPKSF
DGNLLNRGHAIKKASPCJADKISKPOELSSDLNVGDTSONSCVDCSVTQSNKVS
TPPEESQNSDTPPRDLPLDEKHVTSFHPGPNAPIDPDLSEGNSSDINQTKRTY
SLTPSPTTQVETPDLDVHDNTPSLFTPLSNLPHSDSDSDSDSDSDSDSDSDSD
STASATVSAATSTESISTRKVLPMISIARHNTAGTHSGAEKDDVDVSDSPPLPERT
ESFVLASEHNTPVRSSELSQERSEQKSEGLITSENEKCDHPAGGITHYEMCIECP
PTFSDKREQISENPTEADIGFNGRCGPKGRDPPSEWT"

BASE COUNT 1004 a 556 c 577 g 801 t
ORIGIN

alignment_scores:

Quality: 1157.00 Length: 783
Ratio: 2.400 Gaps: 25
Percent Similarity: 61.558 Percent Identity: 37.037

alignment_block:

US-09-600-358A-4 x HUMPTPG1 ..

Align seg 1/1 to: HUMPTPG1 from: 1 to: 2938

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
20 ATGAGCAAGTGGAGATCTTCAGGAAATTCACAGAGGTCCAGGCAT 69
|||||
17 sLysIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||
70 GAGAGTCTCGACCAATGGGAGGACAACTTCGCCGGGACTTCATGC 119
|||||
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
|||||
120 GGTTAAGAAGATTCTCTCCAAATATAGACAGAAAGATATATATCCACA 169
|||||
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspII 63
|||||
170 GCCACTGGAGAAAAGAAAATGTTAAAAAGAACAGATACAGAGCAT 219
|||||
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||
220 ACTGCCATTTGATCAGCCGAGTAAATTGACATTAAAGACTCCTCTCAC 269
|||||

```
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 AAGATTTCAGACTATATCAATGCAAAATTTTATAAGGCGCTCTATGGGCCA 319
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 AAAGCATATGTAGCAACTCAAGGACCTTTAGCAAAATACAGTAATAGATT 369
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 eTrpArgMetIleTrpGluTyrSerValLeuIleIleValMetAlaCysM 130
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 TTGGAGATGGTATGGAGATATAATGTGTGATCATTTGTAATGGCCGTGCC 419
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 etGluTyrGluMetGlyLysLysCysGluArgTyrIleAlaGluPro 146
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 GAGAAATTTGAGATGGGAAGAAATGTCAGCGCTATTGGCCCTTTGTAT 469
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 GlyCysMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG1 163
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 GGAGAAGACCCCATAGCTTTGGACCAATTTAAATTTCTTGTGAGGATGA 519
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsnS 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 ACAGCAAGACAGACTACTCTATCAGACACTTACTTGAATTTCAA 569
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 ATGAATCTCTAGCTCTATCAGTTCTTCAATGTGTAAGTGGCCAGACCAT 619
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
197 AspValProSerSerIleAspProIleLeuGluLeuIleTrpAspValAr 213
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
620 GATGTTCTCTCATCATTTGATTTCTATTCTGGACATGATAAGCTTAATGAG 669
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 qCysTyrGlnGluAspSerValProIleCysIleHisCysSerAlaG 230
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
670 GAAATATCAAGACATGAAGATGTCTTATTGTATTCAATTCAGTCAGCGAG 719
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
720 GCTGTGGAACAACAGGTGGCCATTTGTGCCATA...GATTATACGTGGAAT 766
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuI1 263
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
767 TTACTAAAGCTGGGAAATACCAGAGGAATTAATGTATTATTATTAAAT 816
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGlnT 280
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
817 ACAAGAAATGAGACACAAAGGCAATTCGCAGTACAAACAAGAGGCAAT 866
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
867 ATGAACCTGTTTCATAGAGCTATTGCCCAACTGTTTGAAACACAGCTACAA 916
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 Val.....IleArgAsp.....LysHisSe 303
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
917 CTATATGAAATTCATGGAGCTCAGAAAATTCGCTGATGCAGTGAATGAAAT 966
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
303 rGlyThrGluSerGlnAlaLysHisCysIleProGluLysAsnHisThrL 320
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
967 TACACTGAAACATGTCAGCTCCATAGACCTGAAAAACAA..... 1009
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 euGlnAlaAspSerTyrSerProAsnLeuProLysSerThrThrLysAla 336
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1010 .....GATTCTCTCTCCAAAACCCACCAAGGCCCGCAGT..... 1045
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
337 AlaLysMetMetAsnGlnArgThrLysMetGluIleLysGlu..... 351
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1046 .....TGCCCTTGTGAAGGGGATGCTTAAAGAAATACTGCAGCCACC 1089
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 .....SerSerSerP 355
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1090 GGAACCTCATCCAGTGGCCACCCATCTTGACACCTTCTCCCCCTTCAGCTT 1139
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 heAspPheArgThrSerGluIleSerAlaLysGluLeu..... 368
```

```
1140 TTCCAACAGCTACTACTGTGTGGCAGGACAATGATAGATACCATCCAAG 1189
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 ...ValLeuHisProAlaLysSerSerThrSerPheAspPheLeuGluLe 384
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1190 CCACTGTTGCATATGGTTTCATCAGAACACATTCAGCAGACCTCAACAG 1239
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 uAsnTyrSer..... 387
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1240 AAACATATAGTAAATCAACAGAACTCCAGGGAAAAATGAATCAACAATTG 1289
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 .....PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1290 AACAGATAGATAAAAAATTCGACGAAATTTAAGTTTGTAGATTAAAGAAG 1339
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
403 PheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuG1 419
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1340 GTCCCTCTCCAAGAGGACCA.....AAAAGTTTGTATGGGAA 1377
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
419 ySerLeuLeuPheGluGly.....C 426
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1378 CACACTTTTGAATAGGGGACATGCATTAATAATTAATCTGCTTCACCTT 1427
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
426 ys.....SerAsnSerLysPro.....ValAsn 433
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1428 GTATAGCTGATAAAATCTTAAGCCACAGGAATTAAGTTCAGATCTAAAT 1477
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1478 GTCGGTGATACTTCCCAAGAAATCTTGTGTGAGCTGCAGTGTAAACACAATC 1527
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 r.....ThrProPheGluLeuIleGlnGlnArgGluPhrL 462
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1528 AAACAAAGTTTCAGTTACTCCACCAGAGAATCCAGAAATTCACACACAC 1577
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 ys.....GluValAspSerLysGluAsnPheSerTyr 472
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1578 CTCCAAGGCCAGACCGCTTGCTCTTGATGAGAAAGGACATGTAACGTGG 1627
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1628 TCATTTTCATGGACCTGAAATGCC.....ATACCCATACCTGA 1665
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
489 sValMetHisValSerSerAlaGluLeuAsnTyr..... 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1666 TTTATCTGAAGGCAATTCCTCAGATATCACTATCAAACTAGGAAAACGTG 1715
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 ..SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 516
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1716 TGACTTTAACACCAAGTCCTACACACAAGTTGAAACACCTGATCTTGTG 1765
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVa 533
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1766 GATCATGATAACACTTCA.....CCACTCT 1791
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
533 lGluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysM 550
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1792 CAGAACACCCCTCAGTTTACTAATCCACTCACTCTGATGACTCA.... 1837
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1838 .....GACTCAGATGAAAGAAACTCTGATGCTGTGCTGACCCAGAAT 1879
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1880 AAAACTAATATTTCAACAGCAAGTGCACAGTTCCTGCTGCCTAGTAC 1929
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
582 sSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnG 599
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1930 TGAAGCATTTCTACTAGGAAGATTTGCCAATGTCCATTCCTAGACATA 1979
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
599 lnGluSerAlaValLeuAlaThrAlaProArg.....IleAsp 611
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


1290 ACACGATAGATAAAAATTTGGACAGAAATTTAAAGTTTGTGAGATTAGAAG 1339
403 PheProIleValGlycyluProLeuGlnLysHisGlnSerLeuAspLeuG1 419
1340 GTCCCTCTCCAGAGGCCA.....AAAGTTTGTAGTGGA 1377
419 ySerLeuLeuPheGluGly.....ValAsn 433
1378 CACACTTTTGAATAGGGACATGCAATTAATAATTAATCTGCTTCACCTT 1427
426 ys.....SerAsnSerLysPro.....ValAsn 433
1428 GTATAGCTGATAAAATCTCTAAGCCACAGGAATTAAGTTTCAGATCTTAAT 1477
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
1478 GTCGGTGATACTTCCAGAAATCTTGTGTGGACGTCAGTGTAAACACAATC 1527
450 r.....ThrProPheGluLeuIleGlnArgGluThrL 462
1528 AAACAAAGTTTCAGTTACTCCACCAAGAAGATCCACAGAAATTCAGACAC 1577
462 ys.....GluValAspSerLysGluAsnPheSerTyr 472
1578 CTCCAAGGCCAGACGCTTGCTCTGTAGTGAAGAGCAATGTAACTGGTGG 1627
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
1628 TCATTTTCATGGACCTCAAAATGCC.....ATACCCATACCTGA 1665
489 sValMetHisValSerSerAlaGluLeuAsnTyr..... 500
1666 TTTATCTGAAGCAATTCCTCAGATATCAACTATCAAACTAGGAAAACTG 1715
501 ..SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 516
1716 TGAGTTTAACACCAAGTCTCTAACACAAAGTTGAAACACCTGATCTGTG 1765
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVa 533
1766 GATCATGATAACACTCA.....CCACTCTT 1791
533 lGluAsnProTyrPheSerSerTrpProProSerGlyThrSerSerLysM 550
1792 CAGAACACCCCTCAGTTTCTAATCCACTTCACTCTGATGACTCA... 1837
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565
1838GACTCAGATGAAGAAACTCTGATGTGTGTGACCCAGCA 1879
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
1880 AAAACTAATATTTCAACGCAAGTGCCACAGTTTCTGCTGCCACTAGTAC 1929
582 sSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnG 599
1930 TGAAGACATTTCTACTAGGAAAGTATTGCCAATGTCCATGTCTGACATA 1979
599 lGluSerAlaValLeuAlaThrAlaProArg.....IleAsp 611
1980 ATATAGCAGGAACACACATTCAGTGCTGCTGAAAAGATGTTGATGTAGT 2029
612 AspGluIleProProLeuProValArgThrProGluSerPheIleVa 628
2030 GAAGATTACCTCCCTCCCTACCTGGAAGAACTCTCTGATGTTGTGTT 2079
628 lValGluGluAlaGlyPheSerProAsnValProLysSerLeuSers 645
2080 ACGCAAGTGAA.....CATAAATACACTT..... 2101
645 eAlaValLysValLysIleGlyThrSerLeuGluTrpGly..... 658
2102GTAAGATCCGAATCGAGTGAACCTCAA 2128

388PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
 :::|||||::: :: :::~::~:~::~:~:: ~||


```
1020 .....GATTCCTCCTCCAAACACCAAGGACCCGAGT..... 1055
337 AlaLysMetMetAsnGlnGlnArgThrLysMetGluLeuLysGlu..... 351
1056 .....TGCCCTGTGTAAGGGGATGCTAAAGAGAATACTGACGCCACC 1099
352 .....SerSerSerp 355
1100 GGAACCTCATCCAGTCCACCCATCTTCACACCTTCCTCCCTTCAGCTT 1149
355 heAspPheArgThrSerGluLeuSerAlaLysGluGluLeu..... 368
1150 TTCCAACAGTCACTACTGTGGCAGGACAATGATAGATACCATCCAAAG 1199
369 ...ValLeuHisProAlaLysSerThrSerPheAspPheLeuGluLe 384
1200 CCAGTGTTCATATGTTTCATCAGAACAACTTCAGCAGACCTCAACAG 1249
384 uAsnTyrSer..... 387
1250 AAACATATAGTAATCAACAGAACTTCACGGGAAAAATGAATCAACAATTG 1299
388 .....PheAspLysAsnAlaAspThrThrMetLysTrpGlnThrLysAla 402
1300 AACAGATAGATAAAAAATTGGAACGAAATTTAAGTTTGAGATTAAGAAG 1349
403 PheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAspLeuG 419
1350 GTCCCTCTCCCAAGAGGGACCA.....AAAAGTTTGTGGGAA 1387
419 ySerLeuLeuPheGluGly.....C 426
1388 CACACTTTGTAATAGGGGACATGCAATTAATAATCTCTCTCACCTT 1437
426 ys.....SerAsnSerLysPro.....Valasn 433
1438 GTATAGCTGATAAAATCTCTAAGCCACAGGAATTAAGTTACAGATCTAA 1487
434 AlaAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSe 450
1488 GTCGGTGATACTCCCAAGATCTGTGTGGAGTCGAGTGAACACAAATC 1537
450 r.....ThrProPheGluLeuGlnGlnArgGluThrL 462
1538 AAACAAAGTTTCAGTTACTCCACAGAGAAATCCACAGAAATTCACACAC 1587
462 ys.....GluValAspSerLysLysGluAsnPheSerTyr 472
1588 CTCCAAGGCCAGACCGCTTGCCTCTTGATGAGAAGGACATGTAACTGG 1637
473 LeuGluSerGlnProHisAspSerCysPheValGluMetGlnAlaGlnLy 489
1638 TCATTTTCATGACCTGAAATGGCC.....ATACCCATACCTGA 1675
489 sValMetHisValSerSerAlaGluLeuAsnTyr..... 500
1676 TTTATCTGAAGCGCAATCTCTCAGATATCACTATCAAACTAGGAAACTG 1725
501 ..SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVal 516
1726 TGAGTTTAAACCAAGCTCTACACACAAAGTTGAACACCTGATCTTGTG 1775
517 LysHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuVa 533
1776 GATCATGATACACTTCA.....CCACTCTT 1801
533 lGluAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysM 550
1802 CAGAACACCCCTCAGTTTACTAATTCACCTTCACTCTGATGACTCA.... 1847
550 etSerLeuAspLeuProGluLys...GlnAspGlyThrValPheProSer 565
```

```
1848 .....GACTCAGATCAAAAGAACTCTGATGCTGTGTGACCCAGAAT 1889
566 SerLeuLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHi 582
1890 AAAACTAATATTCAACAGCAAGTGCACAGTTTCTGCTGCCACTAGTAC 1939
582 sSerSerSerLeuAsnSerProThrAsnIleSerSerLeuLeuLeuAsnG 599
1940 TGAAGCATTTCTACTAGGAAGATATCCCAATGTCATTCCTAGACATA 1989
599 lNgluSerAlaValLeuAlaThrAlaProArg.....Ileasp 611
1990 ATATAGCAGGAACAACACATTCAGGTGCTGAAAAAGATGTTGATGTTAGT 2039
612 AspGluIleProProLeuProValArgThrProGluSerPheIleVa 628
2040 GAAGATTACCTCTCTCCCTACCTGAAAGAACTCCTGAATCGTTTGTTG 2089
628 lValGluGluAlaGlyGluPheSerProAsnValProLysSerLeuSers 645
2090 AGCAAGTGAA.....CATAATACACT..... 2111
645 erAlaValLysValLysIleGlyThrSerLeuGluTrpGly..... 658
2112 .....GTAAGATCGGAATGGAGTGAACCTCAA 2138
659 .....GlyThrSerGluProLysLysPheAspSerValIleLeuAr 673
2139 AGTCAGGAACGATCTGAACAAAAAAGTCTCAAGC.....TTGAT 2179
673 gProSerLysSerValLysLeuArgSerProLysSerGlyLysAsnPhe 689
2180 AACCTCTCAAAATGAGAAATGTGATCATCCACGGGAGGTATTCACTAT 2228
```

seq_name: gb_ro:MMPTPPES

seq_documentation_block:

LOCUS MMPTPPES 2676 bp DNA linear ROD 18-SEP-1995
DEFINITION M.musculus mptp-pest gene.
ACCESSION X86781
VERSION X86781.1 GI:804997
KEYWORDS protein-tyrosine-phosphatase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .2676

/organism="Mus musculus"

/strain="BALB/c"

/db_xref="taxon:10090"

/map="region A3-B"

/clone_lib="Lambda zap II"

/dev_stage="17.5 days"

1. .14

/note="synthetic linker"

330. .2657

/gene="mptp-pest"

330. .2657

/gene="mptp-pest"

```
/EC_number="3.1.3.48"
/codon_start=1
/product="protein-tyrosine-phosphatase"
/protein_id="CAA60477.1"
/db_xref="GI:804998"
/db_xref="MGI:104673"
/db_xref="SWISS-PROT:P35831"
/translation="MEQVEILRRFQVQAMKSPDHNGEDNFARDFMRLRLSTKYRT
EKIYPTAGTEENVKKRYKDIILFDHSRVKLTLPKPSODSDVINANFIKGVGPKA
YVATQPTGRNEVDFWMIWEYNVMVMACREPMGRKKCERYVPLXGEDIITFAPF
KISENEQARDIYFIRLULLEFQNESRLIOFHVWNPDPVPSFDSILDMILMRK
YOEHEVDVPCIHCSAGCGRTGALCAIDYTNLLKAGKIPFEFNFLIENMRTQRHSA
VQTEKSLYCHRAQLEKQIQLEIYHGAQKIRIDNEIITGTWVSSIDSSEKQSDPPP
KPPRTSLVSGDAKEELQPEHPHPVPIITPPSPAPFTVTTWQSDSDRYHPKPV
HMASPEOHADLNRSYKSDQWGSSEAIIEHDKLERNLSFEIKKVPLOEGPKSFD
GNTLLNRGHAIKIKSASSVVDRTSKPOELSAGALKVDDVDSQNSCADCSAHSRAAE
SSESQNSHRTPPRDLQPKGHVWLSLHGPENATVPDPSPDGKSPDNHSQITKTV
SSTPNTEAEHAHDLTEHNSPLKAPLSFTNPLHSDMDHSDGSSDGAVTRNKTSI
STASATVSPASAESACHRRVLPMSIARQEVAGTPHSGAERKADVSESPPLPERTP
ESFLVADMPVPEHELPNQWSEIORESEGLTTSQNEKHDAAGIHTHEASADSPPAFSD
KKDQITKSPAETVDIGFGRGCKPKPREPSEWT"
BASE COUNT      774 a  632 c  682 g  588 t
ORIGIN

alignment_scores:
  Quality: 1130.50      Length: 751
  Ratio: 2.490          Gaps: 27
  Percent Similarity: 60.453  Percent Identity: 37.017

alignment_block:
US-09-600-358A-4 x MMMPPTPES ..
Align seg 1/1 to: MMMPPTPES from: 1 to: 2676

1  MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
330  ATGAGCAAGTGGAGATCTCGAGGAGGTTCCAGAGGCTCAGAGGCTCAGGCGCAT 379
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
17  sLysIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
380  GAGAGTCCGGATCACAAATGGGAGGACAACTTCGCCGGGAGCTTCAATGC 429
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
30  ysLeuLysArgGlnSerThrLysLysAlaAspLysThrLysThrProThr 46
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
430  GATTGAGAAGATTCTCTACCAATATAGAACAGAAAGATTATCCACACA 479
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
47  ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspII 63
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
480  GCCACTGGAGAAAAGAGAAATGTTAAAAGAACAGATATAAGGACAT 529
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
63  eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
530  ACTGCCATTTGATCACAGCGAGTTAAGTTGACTTTGAAGACTCCATCCC 579
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
80  luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
580  AAGATTGAGATTATATCAATGCAAAATTTTATTAAGGTGTGTATGGGCCA 629
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
97  LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
630  AAAGCATATGTGGCAACCCCAAGGCGCTTCGGGAATACAGTCATAGACTT 679
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
113  eTTPArgMetIleTrpGluTyrSerValLeuIleIleValMetAlaCysM 130
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
680  CTGGAGGATGATATGGGAGTATATATGTTGATGATCGTGTATGGCCGTGC 729
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
130  etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
730  GAGAAATTTGAGATGGGAGGAGAAAGTGTGAGCGCTACTGGCCCTTGTAT 779
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
147  GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGl 163
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```



```

seq_name: gb_pat:AX306312

seq_documentation_block:
LOCUS      AX306312               2983 bp    DNA        linear    PAT 11-DEC-2001
DEFINITION Sequence 1063 from Patent WO0188188.
ACCESSION  AX306312
VERSION    AX306312.1 GI:17645552
KEYWORDS
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (sites)
AUTHORS   Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE     Method for examining ischemic conditions
JOURNAL   Patent: WO 0188188-A 1063 22-NOV-2001;
          School Juridical Person Nihon University (JP)
FEATURES   Location/Qualifiers
            source          1..2983
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
BASE COUNT  923 a   634 c   681 g   745 t
ORIGIN
alignment_scores:
      Quality: 1104.50      Length: 743
      Ratio: 2.493          Gaps: 22
      Percent Similarity: 59.623      Percent Identity: 35.935

alignment_block:
US-09-600-358A-4 x AX306312

Align seg 1/1 to: AX306312 from: 1 to: 2983

1 MetAspGlnArGgLuIlleGluGlnLysPheLeuAspGluAlaGlnSerIly 17
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
52 ATGGAGCAAGTGGAGATCCTCCTGAGGAGGTTCATCCAGAGGGTCCAGGCCAT 101
17 sLysileThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
102 GAAGAGTCCGGATCACATGGGAGGAGCAACTTCGCCGGGACATTCATGCG 151
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
152 GATTGAGAAGATTGCTACCAAAATATAGAACAGAAAGATTATCCACA 201
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIl 63
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
202 GCCACTGGAGAAAAGAGAAATGTTTAAAGAACACATATAGGACAT 251
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
252 ACTGCCATTGATCACACCCAGCTTAAGTTGACTTTGAAGACTCCATCCC 301
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
302 AAGATTTCAGATTATATCAATGCAAAATTTTAAAGGGTGTGTATGGGCCA 351
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
352 AAAGCATATGTGGCAACCAAGGCCCTTTCGGGAATACAGATCATAGACTT 401
113 eTrpArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysM 130
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
402 CTGGAGGATGATATGGGAGTATATATGTTGATGATCGTATGGCCTGTC 451
130 etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
452 GAGATTGAGATGGGAGGAAAAGTGTGAGCGCTACTGGCCCTTGTGAT 501
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```


DEFINITION M.musculus mRNA for p19-protein tyrosine phosphatase.

ACCESSION X63440.1 GI:416181

VERSION X63440.1

KEYWORDS p19-PTP gene; protein tyrosine phosphatase.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2983)

AUTHORS Den Hertog, J.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-1991) J. den Hertog, Hubrecht Laboratory,

3584 Ct Utrecht, THE NETHERLANDS

REFERENCE 2 (bases 1 to 2983)

AUTHORS den Hertog, J., Pals, C.E., Jonk, L.J. and Kruijer, W.

TITLE Differential expression of a novel murine non-receptor protein

tyrosine phosphatase during differentiation of p19 embryonal

carcinoma cells

JOURNAL Biochem. Biophys. Res. Commun. 184 (3), 1241-1249 (1992)

FEATURES 92272714

source Location/Qualifiers

1..2983

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="P19 EC/3T3"

/clone_lib="zap/gt11"

52..2373

/gene="p19-ptp"

52..2373

/gene="p19-ptp"

/note="PTP1B related"

/codon_start=1

/product="p19-protein tyrosine phosphatase"

/protein_id="CA45037.1"

/db_xref="GI:416182"

/translation="MEQVELRRFIQVQAKSPDNGEDNFARDPMRLRLSTKYRT

EKIYPTATGEKENYKKNRYKIDLPDHSRVKLTKLTPSDSDYINANIKGVYGPKA

YVATQGFNFTVIDFWRMIWYVNMVACREFEMGRKCKERYWPLYGEDRITFAPF

KISCNEQARTDIFITLLEFQNESRLIYQHYNWPDHPVSSFDSDIMLSLMRK

YQHEDEVCICACGGRTGAIADITWNLLKAGKIPFENFVFLIQEMRTQRHSA

VQTKQELVHRVIAQIFNSYCMFEMHRSVMYMKLPLELWSVPLIARLDLSAKA

AADSLPCRGCGRTATRTSPGATHPDAITFSLPNVTTCRTVTGTQSCCTW

PHOSNTPSTAMINORTNGAKESAIEHDKLERNLSEFIKKVPLQGPCKSDGN

TLLNRGHAIKISASSVVDRTSKPQELSGALKVDVDSQNSCADCSAAHSHRAESS

ESQSNHPTPPDCPLDKKGHTVWSLGNFNPVPSDGGKSPDNHSQTLTKTVSS

TPNSTAEAEADHDLTEHNSPLIKAPLSFTNPLHSDHSDGSDGAVTRNKTIST

ASATVSPASAEACHRRVLPMSIARQEVAGTPHSGAEKADADYSESPPLPTPES

FVLADMPVPEWHLPQEWSEQRESEGLTSGNEKHDAAGIHTTEASADSPFAFSDDK

DQITKSPAEVTDIGFCNGCKGKPGPREPPSEWT"

BASE COUNT 923 a 634 c 681 g 745 t

ORIGIN

alignment_scores:

Quality: 1104.50 Length: 743

Ratio: 2.493 Gaps: 22

Percent Similarity: 59.623 Percent Identity: 35.935

alignment_block:

US-09-600-358a-4 x MMP19PTP ..

Align seg 1/1 to: MMP19PTP from: 1 to: 2983

1 MetAspGlnArgGluLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

52 ATGGAGCAAGTGGAGATCCTCAGGAGGTTCATCCAGAGGGTCCAGGCCAT 101

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

17 slvstlethr.....LysGluGluPheAlaAsnGluPheLeu 30

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

102 GAAGAGTCGGATCACAAATGGGGAGGACAACTTCGCCGGGACTTCATGCG 151

30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

152 GATTGAGAAGATTGCTACCAAAATATACACAGAAAAGATTATCCCA 201

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

47 ThrValAlaGluAsnAlaLysAsnLysLysAsnArgTyrLysAsp 63

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

202 GCCACTGGAGAAAAGAAAAGAAATGTTAAAAGAACAGATATAAGGACAT 251

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuLeuThrSerAsp 80

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

252 ACTGCCATTGATCACACCGCAGTTAAGTTGACTTTGAAGACTCCATCCC 301

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

302 AAGATTGAGATTATATCAATCAATCAATTTTAAAGGGTGTGATGGGCCA 351

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAsp 113

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

352 AAAGCATATGTGGCAACCCAGGGCCTTCCGGAATACAGTCATAGACTT 401

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

113 eTrpArgMetIleTrpGluTyrSerValLeuIleValMetAlaCysM 130

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

402 CTGGAGGATGATATGGGAGTATATGTTGTGATGATCGTATGGCGCTGC 451

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

130 etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

452 GAGAAATTGAGATGGGAAGAAAAGTGTGAGCGCTACTGGCCTTTGTAT 501

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

502 GGAGAAGATCCTATAACATTTGCACCACTTTAAATTTCTGTGAAAATGA 551

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

163 uLysArgLysSerAspTyrIleArgThrLeuLysValLysPheAsnS 180

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

552 ACAAGCAAGAACCGACTACTTCTATCCGAAACACTTTTACTGAAATTTCAA 601

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

602 ATGAATCCCGTCGGCTCTATCAGTTTCTATACGTCAGTGGCCAGACCAT 651

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

197 AspValProSerSerIleAspProIleLeuLeuLeuIleTrpAspVal 213

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

652 GATGTTCTTCCTGCTCATTTGATTCTATTCTGACATGATAAGCTTAATGAG 701

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

213 qCysTyrGlnGluAspSerValProIleCysIleHisCysSerAlaG 230

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

702 GAATACCAAGAACATGAAAGATGGCTATTGTTGTTGTTGTTGTTGTTG 751

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

230 lCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

752 GCTGGGAGAACAGGTGCTATTGTCGATA...GATTACACGTGGAAC 798

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

247 LeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeu 263

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

799 TTAAGTAAAGCAGGAGAAAATTCAGAGAAATTAATGTTATTAATTAAT 848

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 280

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

849 ACAAGAAATGAGACACAGAGGCTCGGACAGTACAAACAAAGAGGAGCAGT 898

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

899 ATGAACTTGTTTCATAGGGCTATTGCTCAACTGTTTGAACACAGCTACAC 948

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

297 ValIle..... 298

|||||

949 TGTATGAAATTCATGGAGCACAGAGATCCGTTGATGTTAATGAAATTAAC 998

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

299ArgAspLysHisSer. 303

|||||

999 ACTGGAACACTATGGTCAGTTCCATTGATAGCAGAGACGACCTCTCCGCCA 1048

|||||

304GlyThr 305

```
1049 AAGCCGCGGGGACTCGAAGTTCCTTGTAGAAAGGGGATGCCAAGGAAGA 1098
|||||
306 GluSerGlnAlaLysHisCysIleProGluLysAsnHisThrLeuGlnAl 322
|||||
1099 AATACTACAGCCACCAACCACTCACCGGTGCCACCAACCATCTCAGCCCAT 1148
|||||
322 aAspSerTyrSerProAsnLeuProLysSerThrThrLysAlaAlaLysM 339
|||||
1149 CACCTCCTTCAGC.....CTTCCACAGTTACCACCTGTGTGCAGGACAG 1192
|||||
339 et.....MetAsn 341
|||
1193 TGACAGGTACCAACCAAGCCAGTGTGCACATGGCCTCACAGAGCAAC 1242
|||||
342 GlnGlnArgThrLysMetGluLysGluSerSerSerPheAspPheAr 358
|||||
1243 ACCAGCGGACCTCAACAGAA.....GCTATGATAAATCAGCG 1280
|||||
358 gThrSerGluIleSerAlaLysGluGluLeuValLeuHisProAlaLys 375
|||||
1281 GACC.....AATGGGCAAAA 1297
|||||
375 erSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysAsn 391
|||||
1298 GTGAATCAGCTATTGACACATA.....GATAAGAAG 1329
|||||
392 AlaAspThrThrMetLysTyrGlnThrLysAlaPheProIle..... 405
|||||
1330 TTAGAGCGCAATTAACTTTTGTAGATTAGAAGTCCCTCTCCAAGAAGG 1379
|||||
405 ..... 405
1380 GCCCAAAAGTTTGTAGGGAACACACTCTTGAATAGGGGACATCGGATTA 1429
|||||
406 .....ValGlyGluProLeuGlnLysHisGln 414
|||||
1430 AAATTAATCTGCTTCATCTCTGTAGTTGACAGAACCTCTAAACACACAG 1479
|||||
415 SerLeuAspLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysPr 431
|||||
1480 GAGTTAAGTCAGGTGCCCTAAAGGTTGATGATATCTCAGAATCTTG 1529
|||||
431 oValAsnAlaAlaGlyArgTyr..... 438
|||||
1530 CGCGGACTGTAGTCGCGCTCATTCACACAGAGCTGTGAGTCGTCAGAGG 1579
|||||
439 .....PheAsnSerLysValProIleThrArgThrLysSerThrPro 452
|||||
1580 AGTCCCAAGCACTACACACACCT...CCACGCCAGACTGTGTCCT 1626
|||||
453 PheGluLeuIleGlnArgGluThrLysGluValAspSerLysGluAs 469
|||||
1627 CTCGATAG...AAAGACAGTAACTGGTGGTCATTCATGACCTGAAAA 1673
|||||
469 nPheSerTyrLeuGluSerGlnProHisAspSerCysPheValGluMetG 486
|||||
1674 TGCCACACCTTACCCGACCTACCT...GACGGCAATCCACAGATAATC 1720
|||||
486 lnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyrSerLeu 502
|||||
1721 ATTCTCAGACTCTGAAACCGTGTGATTCACACCC.....AACTCCACC 1764
|||||
503 ProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnValLysHisH 519
|||||
1765 GCAGAAGAGGAGCCACGATCTT.....ACAGAGCACCA 1799
|||||
519 sAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValGluAsnP 536
|||||
1800 CAACAGCTCC.....CCTCTGTGAAGCTC 1825
|||||
536 ro.....TyrPheSerSerTyrProProSerGly 545
|||||
```

```
1826 CCCTCAGCTTTACCAACCCCTTCTTCACTCTGACGACTGGCACTCAGCGGA 1875
|||||
546 ThrSerSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrVa 562
|||||
1876 GGAGGCTCT.....GATGCTGCTGT 1895
|||||
562 lPheProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrT 579
|||||
1896 GACCAGGAACAAATAGCATTTCAACAGCAAGTGCACACAGTGTCTCCTG 1945
|||||
579 yrAsnSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSer 595
|||||
1946 CCAGTAGTGTGAGAGT...GCTTGCATAGGAGAGTATTGCCGATGCTCC 1992
|||||
596 LeuLeuAsnGlnGluSerAlaValLeuAlaThrAlaProArg..... 609
|||||
1993 ATTGCCAGACAGGAA.....GTAGCAGGCACGCCGATTCAGGTGC 2033
|||||
610 .....IleAspAspGluIleProProLeuProVala 621
|||||
2034 TGAGAAAGATGCTGATGTTAGTGAGGAGTGCCTCCTCTTACCTGAAC 2083
|||||
621 rgThrProGluSerPheIleValValGlu 630
|||||
2084 GAATCCTGAGTCTTTTGTATTAGCAGAT 2112
|||||
seq_name: gb_ro:RATRKPTP
seq_documentation_block:
LOCUS RATRKPTP 1608 bp mRNA linear ROD 08-FEB-1999
DEFINITION Rat mRNA for protein tyrosine phosphatase, complete cds.
ACCESSION D38072
VERSION D38072.1 GI:567262
KEYWORDS protein tyrosine phosphatase.
SOURCE Rattus norvegicus (strain:Sprague-Dawley) kidney cDNA to mRNA,
clone:RKPTP.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus
REFERENCE 1 (bases 1 to 1608)
AUTHORS Moriyama,T.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1994) Toshiki Moriyama, Osaka University Medical
School, 1st Dept. of Medicine; 2-2 Yamadaoka, Suita, Osaka 565,
Japan (Tel:06-879-3632, Fax:06-879-3639)
REFERENCE 2 (bases 1 to 1608)
AUTHORS Moriyama,T., Kawanishi,S., Inoue,T., Imai,E., Kaneko,T., Xia,C.,
Takenaka,M., Noguchi,T., Kamada,T. and Ueda,N.
TITLE cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPTP)
from rat kidney
JOURNAL FEBS Lett. 353 (3), 305-308 (1994)
MEDLINE 95046282
FEATURES
source
1..1608
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/gb_xref="taxon:10116"
/clone="RKPTP"
/tissue_type="kidney"
135..1283
/codon_start=1
/product="protein tyrosine phosphatase"
/protein_id="BAA07266.1"
/db_xref="GI:699627"
/translation="MEQVEILRRFTQRVQAMKSPDHNGEDNPFARDFMLRLRLSTKYRT
EKIYPTATGEKEENVKKNRYKDILFDHSRVKLTLKTPSQSDIYNANFIKGVYGPRA
VYATQGLANTVIDFWIMWYNNYIIVMACREFEMGRKKERYWPLXGDDITTFAPF
KISCENEQARTDYFRTLLLEFQNESRRLYOFHYVNPDPHVPSSFDLDMISLMRK
YOEHEVDPTCHCSAGCGTGAICAIDYTNLLKAGKIPPEENVNLIQEMRTORHSA
VOTKEQYELVHRAIQLEKOLQYIHAOKITDGNISTGNMYSIDSEKQDSDPPP
KPPTRSCSLVEGDAREEILQPPFHPVLPILTPSPPSAFTLHCVARQ"
BASE COUNT 490 a 347 c 410 g 361 t
```

ORIGIN

alignment_scores:
 Quality: 1058.00 Length: 465
 Ratio: 3.112 Gaps: 12
 Percent Similarity: 73.118 Percent Identity: 47.742

alignment_block:
 US-09-600-358a-4 x RATRKPTP ..

Align seg 1/1 to: RATRKPTP from: 1 to: 1608

```

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerly 17
|||||
135 ATGGAGCAAGTGGAGATTCTGAGGAGGTTCAATCCAGAGGGTCCAGGCCAT 184
17 slysylethr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||
185 GAAAGTCCGGACCACAATGGGGAGGACAACCTCGCCGGGACTTCATGC 234
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
|||||
235 GATTGAGAAGATTGTCTACCAATATAGAACAGAAAAGATTATCCACA 284
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspI 63
|||||
285 GCCACTGGAGAAAAGAAATGTTAAAGAACAGATATAAGGACAT 334
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||
335 ACTGCCATTTGATCAACAGCGAGTTAAGTTGACTTTGAAGACACCATCCC 384
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
|||||
385 AAGATTACAGATTATATCAATGCAATTTTATTAAGGGTGTATATGGCGC 434
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||
435 AGAGCATACGTGGCAACCCAGGGCCTTTGGGAATACAGTCATAGACTT 484
113 eTrpArgMetIleThrGluTyrSerValLeuIleValMetAlaCysM 130
|||||
485 CTGGAGGATGATATGGGAGTACAATGTGTCATCATCTTAATGGCCTGC 534
130 etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
|||||
535 CTGAATTTCAGATGGGAAGGAAAAGTGTGAGCGCTATTGGCCTTTGTAT 584
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaG 163
|||||
585 GGAGAAAGATCCTATTAACATTTGCACCATTTAAATTTCTGTGAAAATGA 634
163 uLysArgLysSerAspTyrIleLeuArgThrLeuLysValLysPheAsnS 180
|||||
635 ACAACGAAGACAGACTACTTCATTCGACACACTTTTACTTGNATTCAA 684
180 erGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
|||||
685 ATGAATCCCGTCTGCTATCATAGTTTCACTACGTGAAGTGGCCAGACCAT 734
197 AspValProSerSerIleAspProIleLeuGluLeuIleTrpAspValar 213
|||||
735 GATGTCCCTTCGTCATCTATCAATCTTATCTGGACATGATAAGTCTAATG 784
213 gCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAlaG 230
|||||
785 GAAGTACCAAGACAGAGAGTGTGCTTATTTGTATCCATTCGACGTGCG 834
230 lyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
|||||
835 GCTGTGGAGCGGAGGTGCTATTTGTGCCATC...GACTACACGTGGGAA 881

```

```

247 LeuLeuLysAspGlyIleLeuProGluAsnPheSerValPheSerLeuI 263
|||||
882 TTACTGAAAGCAGGAAATCCAGAGGAATTAATGATTATTAATTAAT 931
263 eArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 280
|||||
932 ACAAGAAATGAGAACACAAAGCACTCTGCAGTACAAAACGAAGGAGCAGT 981
280 yrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMetAsp 296
|||||
982 ATGAACCTGCTCCATAGACTATTGCCCACTGTTTGAAGAAACAGCTACA 1031
297 Val.....IleArgAspLysHis..... 302
|||
1032 CTGTATGAATCCACGGAGCCAGAAATCACTGATGTAATGAATAATAG 1081
303 .....SerGlyThrGluSerGlnAlaLysHisCysIleP 314
|||||
1082 CACTGGAAACATGGTCAGTTCATTGTAGTGAACAAACAAGATTCTCCTC 1131
314 roGluLys..AsnHisThrLeuGlnAlaAspSerTyrSerProAsnLeuPr 330
|||
1132 CTCCAAAGCCACCGAGGACTCGAAGTT...GTCTTTGTAGAGGGGATGCC 1178
330 oLysSerThrThrLysAlaAlaLysMetMetAsnGlnGlnArgThr.... 345
|||||
1179 AAGGAAGAAATCCTCCAGCCACGAGGCTCACCGGTGCCGCCCATCCT 1228
346 .....LysMetGluIleLysGluSerSerSerPheAspPheArgThrSer 360
|||
1229 GACACCGTCGCTCCTTCAGCATTCCTCAACGTTACACTGTGTGGCAAGAC 1278
361 GluIleSerAlaLysGluGluLeuValLeuHis.....ProAlaLys 374
|||
1279 AGTGACAGGTACCAACCAAGCCAGTGTGCACATGGGCTCCCAAGACA 1328
374 sSerSerThrSerPheAspPheLeuGluLeuAsnTyrSerPheAspLysA 391
|||
1329 GCCCCCAACG.....GACCTCAACAGAACTATGATAAGT 1363
391 snAlaAsp.....ThrThrMet 396
|||||
1364 CAGCGGACCTAATGGGGAGAACGGAATCTGCTGTTGAGCACACAGATAAA 1413
397 LysTrp....GlnThrLysAlaPheProIleValGlyGluProLeuGlnLy 412
|||
1414 AAGTAGAACAAAAATTTAAGTTTGAAGATTAAAGAAAGTCCCTCTCCAAGA 1463
412 s...HisGlnSerLeuAspLeuGlySerLeuPheGluGly 425
|||
1464 AGGGCCCAAAAGTTTTGTGGAACACACACTCTTGAATAGGGGA 1506

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 07:00:23 ; Search time 95.32 seconds
(without alignments)
6071.260 Million cell updates/sec

Title: US-09-600-358A-3
Perfect score: 2356
Sequence: 1 tccctcaacctacttataga.....cctagtcgagcacatccgg 2356

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2074.6	88.1	3580	4	US-09-081-345-1
2	216.4	9.2	1529	4	US-08-821-278A-1
3	214.2	9.1	2226	3	US-08-951-260A-1
4	130.6	5.5	2309	1	US-08-036-210-10
5	130.6	5.5	2309	2	US-08-449-609-10
6	130.6	5.5	2692	1	US-08-036-210-14
7	130.6	5.5	2692	2	US-08-449-609-14
8	130.6	5.5	3973	1	US-08-036-210-21
9	130.6	5.5	3973	2	US-08-449-609-21
10	120	5.1	466	4	US-08-821-278A-16
11	116.8	5.0	5117	3	PCT-US95-05512-1
12	116.8	5.0	5117	5	US-08-854-585-1
13	115.4	4.9	3969	1	US-08-241-853-1
14	115.4	4.9	3969	2	US-08-850-917-1
15	106.2	4.5	247	1	US-08-594-031-74
16	105.6	4.5	2173	1	US-08-036-210-6
17	105.6	4.5	2173	2	US-08-449-609-6
18	104.8	4.4	4752	1	US-08-201-697-3
19	104.8	4.4	4799	1	US-08-201-697-6
20	104.8	4.4	4815	1	US-08-201-697-5
21	100.6	4.3	5455	1	US-08-342-930-1
22	98.4	4.2	4651	2	US-08-449-644-4
23	98.4	4.2	4651	2	US-08-087-244A-4
24	90.4	3.8	5690	2	US-08-447-464-2
25	90.4	3.8	5690	2	US-08-716-679-2
26	90	3.8	6000	1	US-08-348-006B-6
27	90	3.8	6000	2	US-08-800-825A-6

28	90	3.8	6000	4	US-09-158-657-6	Sequence 6, Appli
29	90	3.8	6000	5	PCT-US94-10166-6	Sequence 6, Appli
30	88.8	3.8	5679	1	US-08-201-697-1	Sequence 1, Appli
31	87.2	3.7	5769	1	US-08-652-971-1	Sequence 1, Appli
32	87.2	3.7	5769	2	US-08-991-258A-1	Sequence 1, Appli
33	87.2	3.7	5769	2	US-08-769-399-1	Sequence 1, Appli
34	87.2	3.7	5769	3	US-08-991-953A-1	Sequence 1, Appli
35	84	3.6	2351	2	US-08-548-159-6	Sequence 6, Appli
36	80.8	3.4	6924	1	US-08-015-973-2	Sequence 2, Appli
37	80.8	3.4	6924	2	US-08-448-164-2	Sequence 2, Appli
38	78.8	3.3	4338	1	US-08-015-986A-1	Sequence 1, Appli
39	78.8	3.3	4338	2	US-08-446-363-1	Sequence 1, Appli
40	74.6	3.2	963	3	US-08-964-308-1	Sequence 1, Appli
41	74.6	3.2	963	3	US-08-964-308-7	Sequence 7, Appli
42	74.6	3.2	963	3	US-08-964-313-1	Sequence 1, Appli
43	74.6	3.2	963	3	US-08-964-313-7	Sequence 7, Appli
44	74.6	3.2	963	4	US-09-069-138-1	Sequence 1, Appli
45	74.6	3.2	963	4	US-09-069-138-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-081-345-1
; Sequence 1, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 673510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-081-345-1

Query Match 88.1%; Score 2074.6; DB 4; Length 3580;
Best Local Similarity 99.7%; Pred. No. 0;

Matches 2090; Conservative 0; Mismatches 4; Indels 3; Gaps 1;			
QY	1	tcctcaacctacttatagaactatttttctctctgagcattggaccacaaagagaatc	60
Db	12	TCCTCAACCTACTTATAGACTATTTTCTGCTCGACGATGGACCAAGAGAAATTC	71
QY	61	tgcaagttctctggtgagcccaagcaagaataattactaaagaggtttgccaatg	120
Db	72	TGCAGAGTCTCTGGATGAGGCCCAAGCAAGAAATTTACTAAAGAGGAGTTTGCCCAATG	131
QY	121	aattctgaagctgaaagcgcaatctaccaagtacaaagcagacaaacccatctctacaa	180
Db	132	AATTTCTGAAGCTGAAAAGCAATCTACCAAGTACAAGCAGACAAAACCTATCTCTACAA	191
QY	181	ctgtgctgagaatgccagaataatacaagaaaaacagataataaggatatatttggccctatg	240
Db	192	CTGTGCTGAGAGCCCAAGAAATATCAAGAAAACAGATATAGAGATATTTTGCCCTATG	251
QY	241	attcatagccgggtagaactatccctgataacctctctgatgagattccagctacatcaatg	300
Db	252	ATTATAGCCGGGTAGAACTATCCCTGATTAACCTCTGATGAGGATTCAGCTACATCAATG	311
QY	301	ccaacttcattaaggagtttatggaccacaaagcttatattggccacccagggttcctttat	360
Db	312	CCAACCTTCATTAAAGGAGTTTATGSCCCCAAGGCTTATATTGCCCACCCAGGGTCTCTTAT	371
QY	361	ctacaacctctcgactctgagagatgatttgggaatatagtgctcttcatctgtta	420
Db	372	CTACAACCTCTCGACTCTTGAGGATGATTTGGGAATATAGTGCTTATCATTTGTTA	431
QY	421	tggcatgcatggaatgaaatgggaagaaagtgtgagcgctactgggctgagccag	480
Db	432	TGCGATGCATGGAGTATGAATGGGAAGAAAGTGTGAGCGCTACTGGCTGAGCCAG	491
QY	481	gagagatgcagctggaatttggcccttctctgtatctctgtgaagctgaaaaagaaat	540
Db	492	GAGAGATGCGAGCTGGAATTTGGCCCTTCTCTGTATCTCTGTAAGCTGAAAAAGAAAT	551
QY	541	ctgattataatcaggactctaaagttaagttaagttaagtgaactcgaactctatcacc	600
Db	552	CTGATTATATAATCAGACTCTAAAAGTTAAGTTCAATAGTGAACCTCGAACCTATCTACC	611
QY	601	agttcattacaagaattggccagaccatgatgaccttcatctatagaccctattcttg	660
Db	612	AGTTTCAATACAGAATTGSCCAGACCATGATGTACCTTCATCTATAGACCTTATCTTG	671
QY	661	agctatctgggagtacgtgtgtacaaagagatgacaggtgtcccatatgcaattcaet	720
Db	672	AGCTCATCTGGGATGACGTTGTTACCAAGAGGATGACAGTGTTCCTCATATGCAATCACT	731
QY	721	gcagtgctggctggaagactggttatttggctattgtgctattgtgattatacagatgt	780
Db	732	GCAGTGTGCTGTGGAAGACTGGGTGTTATTTGTGCTA---TGATATATACATGGATGT	788
QY	781	tgtctaaagatgggataaattccctgagaacttcagtttttcagtttgcgcgggaatgc	840
Db	789	TGCTAAAAGATGGATAATTCCTTGAGAACTTCAGTGTGTTTTCAGTTTGATCGCGGAAATGC	848
QY	841	ggacacagagcccttcattagtttcaacgcaggacaaataatgaaactggctacaatgctg	900
Db	849	GGACACAGAGCCCTTCATTAGTTTAAACCGCAGGAACAATATGAACCTGGTCTACAATGCTG	908
QY	901	tattagaactatttaagagacagatggtattatcagagataaacattctggacagaga	960
Db	909	TATTAGAACTATTTTAGAGACAGATGATGTTATCAGAGATAAACATCTCTGGAACAGAGA	968
QY	961	gtcaagcaagcattgtattccttgaaaaaatcacactctcccaagcagactcttattctc	1020
Db	969	GTCAAAGCAAGCATTGTATTCTTGAGAAAAATCACACTCTCCAAAGCAGACTCTTATCTC	1028
QY	1021	ctaatttaccaaaaagtaccacaaagcagcaaaaaatgatgaaccacaaagacaaaaa	1080
Db	1029	CTAATTTACCAAAAAGTACCACAAAAGCAGAAAAATGATGAACCAACCAAGGACAAAAA	1088

QY	1081	tgaaatcaaaatacttctctttagacttctgaaataagtgcacaaagaag	1140
Db	1089	TGGAATCAAAAGAACTTCTTCTTTGACTTTTAGGACTTCTGAAATPAAGTGCAAAAGAAG	1148
QY	1141	agctagttttgcaccctgctaaatcaagcacttcttttagacttctggagctcaaatata	1200
Db	1149	AGCTAGTTTGCACCCTGCTAAATCAAGCACCTCTTTTGACTTCTCTGGAGCTAAATATACA	1208
QY	1201	gttttgacaaaaatgctgacacaaacctgaaatggcagacaaaggcatttccaatagtgc	1260
Db	1209	GTTTTGACAAAAATGCTGACAAACCATGAAATGGCAGACAAAGGCATTTTCCAATAGTTG	1268
QY	1261	ggagcctctcagaagcatcaaatgttgatttgggtcctcttcttgggttggaggtgctt	1320
Db	1269	GGGAGCCTCTCAGAAAGCATCAAAAGTTTGGATTGGGCTCTCTTTTGTGGGGATGTT	1328
QY	1321	ctaaattcaaaactgtaaatgcagcaggaagatatatttaattcaaaagtgccaataaacac	1380
Db	1329	CTAATTTCTAAACCTGTAAATGCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACAC	1388
QY	1381	ggacaaaaatcaactccttttgaattgatcacagcagagagaaaccaaagggtggacagca	1440
Db	1389	GGACCAAAATCAACTCCTTTTGAATTGATACGACAGAGAAACCAAGAGGTGGACAGCA	1448
QY	1441	aggaataactttcttatttgggaatctcaaccacatgattcttgtttttagagatgcagg	1500
Db	1449	AGGAAAACCTTTCTTATTTGGAAATCTCAACCACATGATTTCTTTTGTAGAGATGCAGG	1508
QY	1501	ctcaaaagaataatgcatgtttcttcagcagaactgaattattcaactgccaatagactcta	1560
Db	1509	CTCAAAAGTAATGCATGTTTCTTTCAGCAGAACTGAATATTATTCACATGCCATAGACTCTA	1568
QY	1561	aacacaaatacgttaatgctctaatgtaaagcacaatgactctagctctcttgggtgat	1620
Db	1569	AAACACAAATACGTAATGCTCTAATGTAAAGCACCATGACTCTAGTGTCTTTGGTGTAT	1628
QY	1621	attctcacataccttagtggaaaaatccctatttttcatctgctcccaagtgggtacca	1680
Db	1629	ATTCCTTACATACCTTTAGTGGAAAAATCCTTATTTTTCATCATGCGCCTCCCAAGTGGTACCA	1688
QY	1681	gttctaagatgtctcttatttaccttgcagaagcaagatggaactgttttccctctctc	1740
Db	1689	GTTCTAAGATGCTCTTGATTTTACCTTGAGAAGCAAGATGGAACGTGTTTTCTCTCTCTC	1748
QY	1741	tgttgccaacatctctacatccctcttcttattacaattcaacttcttcttctac	1800
Db	1749	TGTTGCCAACATCCTCTACATCCCTCTTCTCTTATTACAAATTCACATGATTTCTTATCAC	1808
QY	1801	tgaattctccaaccaaataattctctcactattgaaccaggagtcagctgtactagcaactg	1860
Db	1809	TGAATTTCTCAACCAATATTTCTCTACTATTGAACCAAGGATCAGCTGTACTAGCAACTG	1868
QY	1861	ctccaagagtagatgataatccccctccacttctctgtacggacacacctgaatcattta	1920
Db	1869	CTCAAGGATAGATGATGAAATCCCCCTCCACTCTCTGTACGAGACACCTGGAATCATTTA	1928
QY	1921	tgtgtgtgagaagctggagaattctcccaaatgttcccaatcttctatctcagctg	1980
Db	1929	TGTGTGTTGAGGAAGCTGGAGAATTCTCAACAAATGTTCACAAATCCTTATCTCAGCTG	1988
QY	1981	tgaaggtaaaaattggaacatcactggaatggggtggaaactctgaacatctgaacaaagaaattg	2040
Db	1989	TGAAGGTAAAAATGGAACATCACTGGATGGGTGGAACTCTGAACATCTGAACCAAGAAATTTG	2048
QY	2041	atgactctgtacttagaccacaaagcagagtgtaaaactccgaactccataatcag	2097
Db	2049	ATGACTCTGTGATACTTAGACCAAGCAAGAGTGTAAAAACTCCGAGTCTCTAAATCAG	2105

Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-821-278A-1

Query Match 9.2%; Score 216.4; DB 4; Length 1529;
Best Local Similarity 57.4%; Pred. No. 1.9e-50;
Matches 410; Conservative 0; Mismatches 301; Indels 3; Gaps 1;
Qy 200 gaatatcaagaaacagataaggaattttgcccctatgattatagccgggtagaact 259
Db 183 gaacacgaacgaacgcgcacaaagatgctggtagcatatgatgagacaagatcctct 242
Qy 260 atccctgataacccctctgatgaggattccagctacatcaatgccaaacttcattaaaggagt 319
Db 243 ttccctgctccaagagagagagacatgaaattacatcaatgccaaacttcacgggcat 302
Qy 320 ttatggaccacagcttatattgcccacccaggggtccctttattacaacccctccctggactt 379
Db 303 agatggagccagcgcctacattgcagcgaagagccctgctcacacactgttggactt 362
Qy 380 ctggagatgattggaaatagttgctcttattcttattgttgcagtcgcatgagatga 439
Db 363 ctggcgcctggttggaggttggggtcgaagggttaactctgatggcctgtccaagagacaga 422
Qy 440 aatgggaagaaagtgtagcgcctactggctgagccagagagagatgcagctgggaatt 499
Db 423 aaatggagcaggaagtgtaacgctattggccgggagcagagcctctaaagctgg 482
Qy 500 tggcccttctctgtatccctgtgaagctgaaaaagaaatctgtgattatataacaggac 559
Db 483 gcctttctgcatccctgcacaaaggagacaaactgaatgcagacacactcctcaggac 542
Qy 560 tctaaaagttaagttaagtgaactgaactatctaccagttctattacaaagaattg 619
Db 543 cctccaggttacattccagaaggaattccgctctgtgcaccaactacagtatatgtcctg 602
Qy 620 gccagaccatgatgtaccttcatatagaccctattcttgagctcatctgggagtacg 679
Db 603 gccagaccaggggttcccagcaggttctgtatcacattctcaccatggtggaggagcccg 662
Qy 680 ttgttaccagagagatgacaggtttcccatatgcatcattcagctgctgctgtggaag 739
Db 663 ctgctccaaaggcttggaactggaacccctctgtgtccactgagctgctgctgctgagc 722
Qy 740 gactggttattgtgctattgttattatatacatgatgttgcataaagatgggataat 799
Db 723 aacaggttctctgtgcg---tggtgactatgtgagcaggtgtgctgctgaccacagaat 779
Qy 800 tcttgagaactcaggttttctgatttgcgggaatgcgacacagagggccttcatt 859
Db 780 ccttcccaacttcagttccttccaagtggtcctggagatgcggaacagcggcctgcagc 839
Qy 860 agttcaacacaggaacaaatgaactggctcacaatgctattagaaactt 913
Db 840 agtcagacagagagcagctacaggttccctgtaccacacacagtgctcagctatt 893

RESULT 3
US-08-951-260A-1
; Sequence 1, Application US/08951260A
; Patent No. 6004791

GENERAL INFORMATION:
; APPLICANT: Aoki, Naohito
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PROTEIN TYROSINE PHOSPHATASE PTP20
; NUMBER OF INVENTION: AND RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,260A
; FILING DATE: October 16, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,860
; FILING DATE: No 6004791ember 13, 1996
; APPLICATION NUMBER: PCT/1897/00946
; FILING DATE: June 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 227/004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to LRNA
US-08-951-260A-1

Query Match 9.1%; Score 214.2; DB 3; Length 2226;
Best Local Similarity 57.4%; Pred. No. 9.3e-50;
Matches 406; Conservative 0; Mismatches 298; Indels 3; Gaps 1;

Qy 207 aagaaacagatataaggatattttgcccctatgattatagccgggtagaactatccctg 266
Db 199 AAGAAGAACCCCTACAAAGACGTGTACCGTATGATGAGCAGAGATCATCTTTTCCTG 258
Qy 267 ataacctgtgatgaggttcagctacatcaatcccaacttcataaggagattatgga 326
Db 259 CTCAGGAGGAGGACGAGGAGATTACATAATGCCAACTTCACCGGGGCACAGATGGA 318
Qy 327 cccaaggtcttatgtgccaccagggtccctttattctcaaccctctggactcttgagg 386
Db 319 AGCCAGCCCTACATTGGCAGCAGGACCCCTGCCTCACACTCTGTGGACTTCTGGCGC 378
Qy 387 atgattgggaataatagttctcttatctattgttattggtcagtcagtgagatgaatggga 446
Db 379 CTGTTTGGGAGTTTGGAAATCAAGTGTATGTGATGGCTGTCTCAGGAGACAGAAATGGA 438
Qy 447 aagaaaagtgtagcgtactactggctgagccagagagatgagctgggaatttggcct 506
Db 439 CGGAGGAAGTGTGAACGCTACTGGGCCACAGGAGCGGGAGCCTTACAGGCGCGGCTTTC 498
Qy 507 ttctctgtatcctgtgagctgaaaaaggaatactctgattataatcaggactctaaaa 566

Db 499 TGCATCACCTGACAAAGGAGACAGCACTGACTTCGGACATCACTCTCAGGACCCCTCCAG 558
QY 567 gtaagtccaatagtaagaaactcgaactatctaccagtttcattacaagaattggccagac 626
Db 559 GTTACATTCAGAAGAAATCCCGTCTGTGACCACTACATGATGCTTGGCCGGAC 618
QY 627 catgatgaccttcattctatagaccctattcttgagctcatctgggatgtacgtttgttac 686
Db 619 CACGGGTTCCCGAGCAGTTCGCATCACATCTCACCATGGTGAGGAGGCCGCTGCCTC 678
QY 687 caagaggatgacagtggtcccatatgcatcactcagctgctggtggtggaaggactggt 746
Db 679 CAAGGACTTGGACCTGGACCCCTCTGTGTCACCTGCTGCTGCTGCTGCTGCTGCTGCT 738
QY 747 gttattgtgctatgttattatatacatggtgtgttgcctaaagatgggataattcctgag 806
Db 739 GTCTGTGTGTC---TGTGATTACGTGAGGCAGTGTCTTCTGACTCAGACAAATCCACCC 795
QY 807 aacttcagtggtttcagtggttgcctgggaaatcgagacagagggccttcattagttcaa 866
Db 796 AATTTCAGCCTCTTTGAAGTGGTCTGGAGATCGGAAACAGGACCTGCAGCGGTGCAG 855
QY 867 acgaggaacaatatgaactggtctacaatgctgtattagaactatt 913
Db 856 ACAGAGGAGCAGTACAGGTTTCCTGTACACACAGTGGCTCAGCTATT 902

RESULT 4

US-08-036-210-10
; Sequence 10, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,210
; FILING DATE: 23-MAR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2309 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1766
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"

US-08-036-210-10

Query Match 5.5%; Score 130.6; DB 1; Length 2309;
Best Local Similarity 51.1%; Pred. No. 1.5e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtggctgagaatgccaagaatatcaagaaaaacagatatcaaggatatattt 232
Db 184 TTCTTCAACTGATGCTGATCTGCCTTGGAAATAGACAAAAACCCGTTCCCAACATAA 243
QY 233 gccctatgattatagccgggtagaactatccctgataacacctctgatgaggtccagcta 292
Db 244 ACCATATAATAATAACAGAGTAAAGCTAAGCTAGTACGCTAGTGTTCACAGGTTCCGATTA 303
QY 293 catcaatgccaaactcattaaagggagtttatggaccaccaaggttatattgccaccaggg 352
Db 304 TATTAATGCCACCTATATTTCTGGTATTTATGTCCAAATGAATTTATTTGCTACTCAAGG 363
QY 353 tctttatcataccctcctggaactcttgagagatgattgggaatatagtgctcctat 412
Db 364 TCCACTACCAGAACAGTTGGAGATTTTGGAGATGGTGTGGGAAACCCAGACCAAAAC 423
QY 413 cattgttatggcatgcatgagatgataatggaatgggaatggagcgcgtactctgggc 472
Db 424 ATTAGTAATGCTAACACACAGTGTTTGAAAAAGACGATCAGATGCATCAGTATTGGCC 483
QY 473 tgagccaggagagatgcagctggaatttggcccttctctgtatcctgtgaagctgaaaa 532
Db 484 AGAGGACAAACAGCCAGTTACTGTCTTTGGAGATATAGTGTACAAAGCTAATGGAGGA 543
QY 533 aaggaaatctgattataatcagcagctctaaagttaaagtccaatagtgaaactcgaac 592
Db 544 TGTTCAAATAGATTGGCATATCAGGATCTGAAAAATTGAAAGGCATGGGATTCGATGAC 603
QY 593 tatctaccagtttctattacaagaattggccagaccatgatgtaccttctatcatagaccc 652
Db 604 TGTTCACACAGTGTAACTTTACTGCCTGCCGACAGCATGGGTCTCTGAGAACAGCCGCC 663
QY 653 tattcttgagctcatctgggatgtagtctgttaccagaagagatgacagtggttcccatatg 712
Db 664 TCTAATTAATCTTTGTGAAGTTGGTTCGAGCAAGCGGACATGACACACACCTATGAT 723
QY 713 cattcactcagtgctggctggaaggactggtgtatttggctattgttgattatattac 772
Db 724 TGTTCACCTCAGTCTGGAGTTGGAAGAACTGSAGTTTTTATTGCTCTGACCATTTAAC 783
QY 773 a 773
Db 784 A 784

RESULT 5

US-08-449-609-10
; Sequence 10, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2309 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1766
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
; US-08-449-609-10

Query Match 5.5%; Score 130.6; DB 2; Length 2309;
Best Local Similarity 51.1%; Pred. No. 1.5e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtgctgagagatgccagaataatcaagaaaaacagatataaggatatttt 232
Db 184 TTCTTCAACTGATGCTGATCTGCTTGGAAATAGAGCAAAAACCGCTTCCCAACATAAA 243
QY 233 gccctatgattagcgggtagaactatccctgataacctctgagagatccagcta 292
Db 244 ACCATATAATAATAACAGAGTAAGCTGATGCTGAGCTAGCTAGCTTCCAGGTTCCGATTA 303
QY 293 catcaatgccacttcattaaaggagttatggaccacaaagcttatattgccaccagg 352
Db 304 TATTAATGCCAGCTATATTTCTGTTATTTATCTCCAATGAATTTATTCCTACTCAAG 363
QY 353 tctttattacaaacctcctgactctgagagatttgggaatatagtgctcttat 412
Db 364 TCCACTACCAAGAACAGTTGGAGATTTTGGAGAATGGTGGGAAACCAAGCAAAAAC 423
QY 413 cattgtatgcatgcatgagatgataaatgggaagaaaagtgagcgtactgggc 472
Db 424 ATTAGTAACTCTACACAGTGTTTGAAAAGACCGATGATGATGCTAGTATTCAGTATTTGGCC 483
QY 473 tgagccaggagagatgcagctggaatttggcccttctctgtatcctgtgaagctgaaa 532
Db 484 AGAGGACAAACAGCCAGTTACTGCTTTGGAGATATAGTATCAAAAGCTAATGGAGGA 543
QY 533 aaggaatctgattatataatcaggactctaaagttaagttcaatagtgaaactgaac 592
Db 544 TGTTCAAAATAGATGGACTATCAGGATCTGAAAATTTGAAAGGCATGGGGATTCGATGAC 603
QY 593 tatctaccagtttcattacaagaatggccagaccatgatgacaccttcatctatagacc 652
Db 604 TGTTGACAGATGAACCTTACTGCTTGGCCAGAGCATGGGGTTCCCTGAGAACAGCGCCCC 663
QY 653 tattcttgactcactctgggatgtagctgttttaccagagagatgacagtggttcccatag 712
Db 664 TCTAATTCACCTTTGTGAAGTTGTTTCGACCAAGCAGGGGCATGACACACACCTATGAT 723
QY 713 catcactgcagtcgtggtgtggaaggactggttatttctgtctattgttgattatcac 772
|||||
; 724 TGTTCACTGCAGTCTGGAGTTGGAAGAACTGGAGTTTTTTATTGCTCTGGACCACTTTAAC 783
QY 773 a 773
Db 784 A 784

RESULT 6
US-08-036-210-14
; Sequence 14, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,210
; FILING DATE: 23-MAR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2149
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
; US-08-036-210-14

Query Match 5.5%; Score 130.6; DB 1; Length 2692;
Best Local Similarity 51.1%; Pred. No. 1.6e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtgctgagagatgccagaataatcaagaaaaacagatataaggatatttt 232
Db 567 TTCTTCAACTGATGCTGATCTGCTTGGAAATAGAGCAAAAACCGCTTCCCAACATAAA 636
QY 233 gccctatgattagcgggtagaactatccctgataacctctgagagattccagcta 292
Db 627 ACCATATAATAATAACAGAGTAAGCTGATGCTAGCTAGCTTCCAGGTTCCGATTA 686
QY 293 catcaatgccacttcattaaaggagtttattgacccaaggcttatattgccaccagg 352
Db 687 TATTAATGCCAGCTATATTTCTGTTATTTATGTCCAAATGAATTTATTGCTACTCAAGG 746
|||||
QY 353 tctttattacaaacctcctgactctgagagatgatttgggaatatagtgctcttat 412
```

Db 747 TCCACTACCGAAGCAGTGGAGATTTTGGAGAAATGCTGTGGAAACACGAGCAAAAAC 806
Qy 413 catgttattgcatgcatgagtgatgaaatgggaaagtaagtgagcgtactgggc 472
Db 807 ATTAGTAATGCTAACACAGTGTCTTGGAAAAGGCGGATCAGATGCCATCATGATTGGCC 866
Qy 473 tgagccaggagagatgcagctgggaatttggcccttctctgtatctgtgaagctgaaaa 532
Db 867 AGAGCACACAAAGCCAGTACTGCTTTGGAGATATAGTATCAAAAGCTAATGGAGGA 926
Qy 533 aaggaaatctgattatatacaggactctaaaagtttaagttcaatagtgaaactcgaac 592
Db 927 TGTTCAAAATAGATTGACATATCAGGATCTGAAAATTCAAAAGGATGGGATTGCATGAC 986
Qy 593 tatctaccagtttcaatacaagaattggcagaccatgatgtaccttcaatctatagacc 652
Db 987 TGTTCGACAGTGAATCTTACTGCTGCGCCAGAGCATGGGTTCTGTGAGAACAGCGCCCC 1046
Qy 653 tattcttgagctcatctgggatgtacgtgtgttaccgaagagatgacagtggtcccatatg 712
Db 1047 TCTAATTCACITTTGGAAGTTGGTTCGAGCAAGCAGGCGACATGACACCATATGAT 1106
Qy 713 cattcaatgcagtgctggctgtggaagagctgtgttatttggctattgttgattatatac 772
Db 1107 TGTTCACCTGCAGTGTGGAGTTGGAAGAACTGGAGTTTTTATTGCTCTGGACCATTTAAC 1166
Qy 773 a 773
Db 1167 A 1167

RESULT 7

US-08-449-609-14
; Sequence 14, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF INVENTION: PHOSPHATASE
; CORRESPONDENCE ADDRESS: 45
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2149
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-449-609-14

Query Match 5.5%; Score 130.6; DB 2; Length 2692;
Best Local Similarity 51.1%; Pred. No. 1.6e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;
Qy 173 tcctacaactgtgctgagaatgccaaagaatatcaagaaataatcaagaaataaggaataatttt 232
Db 567 TTCTTCACATGATCTGATCTGCCTTGGAAATAGAGCAAAAACCGCTTCCCAACATAAA 626
Qy 233 gccctatgattatagccgggtagaactatccctgataaacctctctgatgaggattccagcta 292
Db 627 ACCATATATATACACAGATTAAGCTGATAGCTACACCTAGCTTCCAGGTTCCGGATTA 686
Qy 293 catcaatgccaaacttcattaaagggagtttatggaccccaaggcttatattgcccacccagg 352
Db 687 TATTAATGCCAGCTATATTTCTGGTTATTTATGTCCAAATGAATTTATTGCTACTCAAGG 746
Qy 353 tcctttatctacaaccctcctggactctctggagatgatttgggaataatagttccttat 412
Db 747 TCCACTACCGAAGCAGTGGAGATTTTGGAGAAATGCTGTGGGAAACACGAGCAAAAAC 806
Qy 413 catgttattgcatgcatgagtgatgaaatgggaaagtaagttgagcgtactgggc 472
Db 807 ATTAGTAATGCTAACACAGTGTCTTGGAAAAGGCGGATCAGATGCCATCATGATTGGCC 866
Qy 473 tgagccaggagagatgcagctgggaatttggcccttctctgtatctgtgaagctgaaaa 532
Db 867 AGAGCACAAAGCCAGTACTGCTTTGGAGATATAGTATCAAAAGCTAATGGAGGA 926
Qy 533 aaggaaatctgattatatacaggactctaaaagtttaagttcaatagtgaaactcgaac 592
Db 927 TGTTCAAAATAGATTGACATATCAGGATCTGAAAATTCAAAAGGATGGGATTGCATGAC 986
Qy 593 tatctaccagtttcaatacaagaattggcagaccatgatgtaccttcaatctatagaccc 652
Db 987 TGTTCGACAGTGAATCTTACTGCTGCGCCAGAGCATGGGTTCTGTGAGAACAGCGCCCC 1046
Qy 653 tattcttgagctcatctgggatgtacgtgtgttaccgaagagatgacagtggttcccatatg 712
Db 1047 TCTAATTCACITTTGGAAGTTGGTTCGAGCAAGCAGGCGACATGACACCATATGAT 1106
Qy 713 cattcaatgcagtgctggctgtggaagagctgtgttatttggctattgttgattatatac 772
Db 1107 TGTTCACCTGCAGTGTGGAGTTGGAAGAACTGGAGTTTTTATTGCTCTGGACCATTTAAC 1166
Qy 773 a 773
Db 1167 A 1167

RESULT 8

US-08-036-210-21
; Sequence 21, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; NUMBER OF INVENTION: PHOSPHATASE
; CORRESPONDENCE ADDRESS: 45
; ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,210
FILING DATE: 23-MAR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-869-8864/9741
TELEFAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3973 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3430
OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-036-210-21

Query Match 5.5%; Score 130.6; DB 1; Length 3973;
Best Local Similarity 51.1%; Pred. No. 2e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtgctgaggaatcccaagaatatcaagaaaaacagatatataagatatattt 232
DB 1848 TTCTTCAACTGATGCTGATCTGCTTGTGAATAGACAAACAAACCGCTTCCCAACATATAA 1907

QY 233 gccctatgattagccggtagaactatccctgataacacctgtagagattccagcta 292
DB 1908 ACATATAATANTACAGAGTAAGCTGATGCTGACCTGAGTGTCCAGGTTCCGGATT 1967

QY 293 catcaatgcacttcatttaaggaggtttatggaccacgaagccttatattgccaccagg 352
DB 1968 TATTAAATGCCAGCTATATTTCTGTTATTTATGTCCAAATGAATTTATTGCTACTCAAG 2027

QY 353 tcttttatcaaacctctcgactctctgagagatttggaatatagttgccttat 412
DB 2028 TCCACTACCAAGAACAGTTGGAGATTTTGGAGAAATGGTGGGAAACACCAAGCAAAAAC 2087

QY 413 cattgttatggcagatgagtagatgaatggaaagaaagtgagcgtactgggc 472
DB 2088 ATTAGTAATGCTACACAGTGTTTGAAAAGGACGATGATGATGATGATGATGATGATG 2147

QY 473 tgagccaggagagatgcagctggaatttgcccttctctgtatccctgtgaagctgaaa 532
DB 2148 AGAGGACAAACAGCCAGTTACTGCTTTGGAGATATAGTATACAAAGCTAATGGAGGA 2207

QY 533 aaggaaatctgattataataatcaggactctaaagttaagttaactagtagtaactgaac 592
DB 2208 TGTTCAAATAGATTGGACTATCAGGGATCTGAAAATTTGAAAGGATGGGGATGTCATGAC 2267

QY 593 tatctaccagtttcattacaagaattggccagaccatgatgtaccitctatctatagacc 652
DB 2268 TGTTCGACAGTGTAACTTTACTGCTGGCCAGGACATGTTGGGTTTCCCTGAGAACAGCGCCC 2327

QY 653 tattcttgagctcatctctgggatgtagcttattaccagaaggatgacagtggttcccatatg 712

DB 2328 TCFAATTCACCTTTGTGAAGTTGGTTCGAGCAAGCGGCACATGACACACACCTATGAT 2387
QY 713 cattcaactgcagtgctgctgaggaactgtgtattatttgctatttattgtattatatac 772
DB 2388 TGTTCACCTGAGTCTGAGTGGAGTGGAGAACTGGAGTTTATTATTGCTCTGGACCATTTAAC 2447

QY 773 a 773
DB 2448 A 2448

RESULT 9
US-08-449-609-21
Sequence 21, Application US/08449609
Patent No. 5952212
GENERAL INFORMATION:
APPLICANT: Moller, Niels P.H.
APPLICANT: Moller, Karin B.
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
PHOSPHATASE
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,609
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,210
FILING DATE: 23-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3973 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3430
OTHER INFORMATION: /note= "N=x-unknown nucleotide"
US-08-449-609-21

Query Match 5.5%; Score 130.6; DB 2; Length 3973;
Best Local Similarity 51.1%; Pred. No. 2e-26;
Matches 307; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 173 tctacaactgtgctgaggaatcccaagaatatcaagaaaaacagatatataagatatattt 232
DB 1848 TTCTTCAACTGATGCTGATCTGCTTGTGAATAGACAAACAAACCGCTTCCCAACATATAA 1907

QY 233 gccctatgattagccggtagaactatccctgataacacctgtagagattccagcta 292


```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gallus domesticus
; STRAIN: DT40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 91..3802
US-08-850-917-1

Query Match 4.9%; Score 115.4; DB 2; Length 3969;
Best Local Similarity 49.3%; Pred. No. 3.2e-22;
Matches 336; Conservative 0; Mismatches 33; Indels 9; Gaps 1;

Qy 164 caaaacctatctcaaaccttggtggaatgccaagaatatacaagaaaacagatataa 223
Db 1881 CACTAAATTTCCATGAGGAGGCCAAGAGGCCATTAATCAGAACAAAAACCGTTACAT 1940

Qy 224 ggatatgttgcctatgattagccgggtagaacctatccctgtataacctctgtgagga 283
Db 1941 TGATATTCTTCATATGATCATAAACCGTGTGAGCTCTCTGAGATTCAGAGACCCAGG 2000

Qy 284 ttccagctacatcaatgccaaacttoattaaggaggtttatggaccgaaggttatatgc 343
Db 2001 ATCAGACTACATCAACGCAAGTTATATTGATGGCTTCAAAGAACCGAGAAAAATACATTGC 2060

Qy 344 caccagggctctttatctacaacccctcctgactcttgaggatgattgggaatatag 403
Db 2061 TGCACAGGCCCCAGGATGAACACCGGATGATTCGGAGATGATCTGGGAACAGAA 2120

Qy 404 tgccttatcatctgttggatgcatgagtgatgaaatggaaagaaaagtgtgagcg 463
Db 2121 AGCAACAATTTATTGTCATGTTACTTCGCTGTGAGGAAGAAACAGGAACAATGTGCCCA 2180

Qy 464 ctactgggctgagccagagagatgagctggaatttggcccttctctgtatcctgtga 523
Db 2181 GTACGGGCCATCAATGGGAGAATGGCTCTGCAACATATGGGGACATAACTGTGAAGATCAA 2240

Qy 524 agctgaaaaaagaaactgattatataatcagactctaaaagttaagttcaatagtga 583
Db 2241 CGAAAGTAAATATGTCAGACTATATATTCAGAAACTGCACATPCACAAATGGAAGAGA 2300

Qy 584 aactgaaac-----tatctaccagttttcattacaagaattggccagaccatgatgt 634
Db 2301 AAGAACATCTGGAAGAGATGTCACCTACATTCAGTTCACCAAGCTGGCCAGACCATGGCGT 2360

Qy 635 accttcactatagaccctattcttgagctcatcttgggagtgacgtttgttaccgaagga 694
Db 2361 TCCCGAGGATCCACATCTCTCTTCAAACTCCGACGAGAGTGAATGCTCTCAGCAACTT 2420

Qy 695 tgacagtttcccatatgcattcactgcagctgctggtgtggaaggagactggttatttg 754
Db 2421 TTTAGTGGCCCAATAGTGGTTCATGTGAGTGTGGAGTGTGGCGCACTGGGACCTATAT 2480

Qy 755 tgcattgttgattatcacatgattgttgcataaagatgggataatctctggaacctcag 814
Db 2481 TGGAAATCAGCTATGTTGGAGGGCTGGATGCAGAGGGCAGAGTGGATGTTATGGGCTA 2540

Qy 815 tgtttcagtttgatcccgga 835
Db 2541 CGTTGTGAAGCTGCGCGGCA 2561
```

```

RESULT 15
US-08-594-031-74
; Sequence 74, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.

Query Match 4.5%; Score 106.2; DB 1; Length 247;
Best Local Similarity 64.4%; Pred. No. 2.9e-20;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 112 ttgcaaatgaatttctgaagctgaaaaggcaatctaccaagtacaaggcagacaaaacct 171
Db 1 TCGCCCGGAGCTTCATGGATTGAGAGATTGTCTACCAATATAGAACACAGAAAGATTT 60

Qy 172 atctacaactgtgctgagaatgccagaataatcaagaaaaacagataaaggatattt 231
Db 61 ATCCACAGCCACTGGAGAAAAAGAGAAATGTTAAAAAGACACAGATATAAGGACATAC 120

Qy 232 tgcctatgattatagccgggtagaacctatccctgataacctctgagagattccagct 291
Db 121 TGCCATTTTGTATGACAGCGGAGTTAAAGTTGACTTTGAAGACTTCCATCCCAAGATTTCAGATT 180

Qy 292 acatcaatgccaaactcattaaaggagtttatgaccaaggcttatattgcccaccagg 351
Db 181 ATATCAATGCAAAATTTTATTAAAGGTGTGTATGGCCCAAAAGCAATATGTGCAACCCCAAG 240

Qy 352 gtccttt 358
Db 241 GGCCTTT 247

Search completed: May 26, 2002, 10:33:59
Job time: 12816 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 06:59:45 ; Search time 5560.98 Seconds
(without alignments)
8865.869 Million cell updates/sec

Title: US-09-600-358A-3
Perfect score: 2356
Sequence: 1 tccctcaacctacttataga.....cctagtcggagcacatcgg 2356

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description

1	2356	100.0	2356	6	AX020263	Sequence
2	2352.8	99.9	2356	6	AF001847	Homo sapi
3	2097	89.0	3058	6	AX020261	Sequence
4	2093.8	88.9	3058	9	AF001846	Homo sapi
5	2074.6	88.1	3580	6	AF149916	Sequence
6	1761	74.7	2877	9	AF150732	Homo sapi
7	1261	53.5	2734	10	MUSPROTYPH	M03888 Mouse prote
8	816.8	34.7	143618	9	AL137856	Human DNA
9	447.2	19.0	1817	9	BC017785	Homo sapi
10	378.2	16.1	3160	9	HUMTPPEST	M93425 Human prote
11	376.6	16.0	2938	9	HUMTPPG1	D13380 Human mRNA
12	376.6	16.0	2938	11	GI8337	X18337 SWISS1985 Er
13	370.2	15.7	2676	10	MMPTPPES	X68781 M.musculus
14	368.6	15.6	1608	10	RATRKPTP	D38072 Rat mRNA fo
15	368	15.6	2983	6	AX306312	Sequence
16	368	15.6	2983	10	MMPL9PTP	X63440 M.musculus
17	243.6	10.3	253	9	HSU69700	U69700 Human prote
18	218	9.3	1529	10	MMPTPHSC	U49853 Mus musculu
19	218	9.3	1559	10	MMU35124	U35124 Mus musculu
20	218	9.3	1568	10	MMU52523	U52523 Mus musculu
21	216.4	9.2	1529	6	AR154418	Sequence
22	216.4	9.2	1643	10	BC008512	Sequence
23	214.2	9.1	2198	10	RNU69673	U69673 Rattus norv
24	214.2	9.1	2226	6	AR095603	X79568 H.sapiens B
25	202.2	8.6	2810	9	HSBDP1	X79568 Human mRNA
26	144.6	6.1	4315	9	HSLCAR	Y00062 Human mRNA
27	144.6	6.1	4597	9	HSLCA	M14342 Mouse 200 k
28	139.6	5.9	4655	10	MUSLY5A	M14343 Mouse Ly-5
29	139.6	5.9	5186	10	MUSLY5B	M92933 Mouse lymph
30	136.4	5.8	5007	10	MUSLY5AA	M10072 Rat leukocy
31	133	5.6	3541	10	RATLICAL	Y00065 Rat mRNA fo
32	133	5.6	4027	10	RNLCAR	AR073853 Sequence
33	130.6	5.5	2309	6	AR073853	Sequence 10
34	130.6	5.5	2309	6	I320377	AR073854 Sequence
35	130.6	5.5	2692	6	AR073854	I32038 Sequence 14
36	130.6	5.5	2692	6	I32038	AR073855 Sequence
37	130.6	5.5	3973	6	AR073855	I32039 Sequence 21
38	130.6	5.5	3973	6	I32039	AR173857 Xenopus 1
39	121	5.1	4920	5	AF173857	AR154432 Sequence
40	120	5.1	4866	6	AR154432	D37781 Human mRNA
41	116.8	5.0	4756	9	HUMTPPS	AR109903 Sequence
42	116.8	5.0	5117	6	AR109903	U10886 Human densi
43	116.8	5.0	5117	9	HSU10886	X58287 Murine mR-P
44	116.2	4.9	4802	10	MMMRPTPU	U40790 Rattus norv
45	115.6	4.9	6314	10	RNU40790	

ALIGNMENTS

RESULT 1	AX020263	Sequence 3 from Patent WO9936548.	2356 bp	DNA	linear	PAT 07-SEP-2000
AX020263	LOCUS	AX020263	Sequence 3 from Patent WO9936548.			
	DEFINITION	AX020263				
	ACCESSION	AX020263				
	VERSION	AX020263.1	GI:10044044			
	KEYWORDS					
	SOURCE	human.				
	ORGANISM	Homo sapiens				
	REFERENCE	1 (bases 1 to 2356)				
	AUTHORS	Roifman,C.M.				
	TITLE	Human lymphoid protein tyrosine phosphatases				
	JOURNAL	Patent: WO 9936548-A 3 22-JUL-1999;				
	FEATURES	HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)				
	source	Location/Qualifiers				
		1..2356				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		42..2120				
		/note="unnamed protein product"				
		/codon_start=-1				
	CDS					

/protein_id="CAC07769.1"
/db_xref="GI:10044045"
/translation="MDQREILQKFLDEAQSKKIKKEEFANFLKLKROSTKYKADKTV
PTTVAENAKNKKKRYKDIPLDYDSRVLSLITDDESSYINANFIMKGVGPXAYIAT
QGPISTLLDFRWIMNEXSVLIIVNACMEYEMGRKKCKERYWAEPGEMLEFGPFSVSC
BAERKRSYIIRTLUKVFNSESTRIIYQPHYKNPDPHVPSSIDILEILWIDVRQYQD
DSVPICHSAGCGRTGVCIAVDYTWMLLDKDGIPENFSVSLIREMRTQPSLVQT
QOVELVYNVLELFLKROMDIVIRDKHSGTESQAKHCIPENKHTLQADSYNLPSKST
KAAKMNOQRKMKELKSSSFDERTSEISAKEELVLPKAKSSTSFDFLELANSFDPKNA
DYTKWQTKAPFIMEPELOKHQSLDLSLLFEGSCNSKPYNAAGRYFNNSKYPITPKS
TPFELIQOREKEVDKSNFSLQPHDSCFVEMQAKVMHVSAAELNYSLPDSDKH
QIRNASNYKHDSALGVYSYIPLVENPFSWPPSPSSKMSLDLDPKQDGTVPFSS
LLPSSLSFYLSHSLNSLIPNLSLNQBSAVLATAPRIDDELPLPVRTPPE
SFIVVEEAGEFSPNVKPSLSSAVKVICTSLWEGTSEPKKFDSDVILRPSKSVKLRS
PKSGKNFSML"

BASE COUNT 750 a 466 c 466 g 674 t

ORIGIN

Query Match		100.0%;	Score 2356;	DB 6;	Length 2356;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2356;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	tcctcacaactactatagactatttttctgtcctgcagcatggaccacaagagaattc	60		
Db	1	TCCCTCAACCTACTTATAGACTATTTTCTGCTCTGCAGCATGGACCAAGAGAAATTC	60		
QY	61	tgcagaagttcttgatgagccgcaaaagcaagaataattactaaagaggttgcgaatg	120		
Db	61	TGCAGAAGTTCCTGGATGAGGCCCAAGCAAGAAATTTACTAAAGAGGAGTTTGCCAAATG	120		
QY	121	aattctgaagctgaaaagcgaatctaccaagtacaagcgacagacaaaacctatcctacaa	180		
Db	121	AATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGCGACAGCAAAACCTATCTACAA	180		
QY	181	ctgtggtgagaatgccagaataatcaagaaaacagatatagaagatatatttgccctatg	240		
Db	181	CTGTGGCTGAGAATGCCAAGAATATACAGAAAACAGATATAGAAGATATTTTGCCCTATG	240		
QY	241	attatagcggatagaactatccctgataacctctgatagattccagctacatcaatg	300		
Db	241	ATTATAGCCGGGTAGAACTATCCCTGATAACTCTGATGAGGATTCAGACTACATCAATG	300		
QY	301	ccaacttcataagggagtttatgaccacaagcttatattgccaccagggtccctttat	360		
Db	301	CCAACCTTCAATTAAGGAGTTTATGACCCCAAGCTTTATATGCCACCCAGGTCCTTTAT	360		
QY	361	ctacaacctcctgactctctggagatgatttgggaatatagtgctccttatcattgtta	420		
Db	361	CTACAACCCCTCTGGACTTCTGGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTA	420		
QY	421	tggcatgcatgagatgaaatggaaagaaaaagtgtagcgctactgtggtgagccag	480		
Db	421	TGGCATGATGAGTATGAATGGGAAAGAAAAGTGTGAGCGCTACTGGGTGAGCCAG	480		
QY	481	gagagatgagctggaaatttggccctttctgtatccctgtgaagctgaaaaaggaat	540		
Db	481	GAGAGATGACGTGGAATTTGGCCCTTTCTCTGATCTCTGTGAAGCTGAAAAGGAAT	540		
QY	541	ctgattataatcaggactctaaaagttaagtccaatagtgaaactcgaaactatatacc	600		
Db	541	CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAAGCTCGAACTATCTACC	600		
QY	601	agtttcattacaagaattggccagaccatgatgtacattcattctatagaccctattcttg	660		
Db	601	AGTTTCATTACAAGAAATGGCCAGACCATGATGATCCTTCATCTATAGACCCCTATTCTTG	660		
QY	661	agctcatctggatgtacgttgttaaccaagagatgacagtggtcccatatgcatcact	720		
Db	661	AGCTCATCTGGGATGTACGTGTGTACCAAGAGATGACAGTGTTCCTCATATGCAITCACT	720		
QY	721	gaagtctggctgtggaaggactggttatttggctatttggattatacaatgatg	780		
Db	721	TGAATTCCTCAACCAATATTTCTCTACTATTGAACAGGAGTCAGCTGTACTAGCAACTG	780		

Db	721	GCAGTGCTGGCTGTGGAAGGACTGGTGTATTATTGTGCTATTGTGTATTATACATGATGT	780		
QY	781	tctcaaaagatgggataattcctcgagaacttcagtggttttcagtttgatccgggaaatgc	840		
Db	781	TCTAAAAGATGGGATAATCTCTGAGAACTTCAGTGTCTTTCAGTTTGATCCGGGAATGC	840		
QY	841	ggacacagagggcttcattagtttcaaacgcaggaacaatatgaactggctcacaatgctg	900		
Db	841	GGACACAGAGGCCTTCATTAGTTCAAACGCGAGAACAAATGAACCTGGTCTACAATGCTG	900		
QY	901	tattagaactatttaagagacagatggtatggttatcagagataaaacattctggaacagaga	960		
Db	901	TATTAGAACTATTTAAGACACAGATGGATGTTATCATAGAGATAAACATTTCTGGAACAGAG	960		
QY	961	gtcagaagcaagcattgtattcctcgagaaaaatcacactctcaagcagactcttattctc	1020		
Db	961	GTCAGAAGCAAGCATGTATTCTCTGAGAAAATCACACTCTCCNAGCAGACTCTTATTCTC	1020		
QY	1021	ctaattttacaaaaagtagccacaaagcagcaaaaaatgatgaaccacaaaaagacaaaa	1080		
Db	1021	CTAATTTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAGGACAAAA	1080		
QY	1081	tggaaatcaaaagaattcttctcttggacttttaggactcttgaataagtgcataagaag	1140		
Db	1081	TGGAATCAAAAGAACTCTTCTCTTTGACTTTTAGGACTTCTGAAATTAAGTGCAAAAAGAG	1140		
QY	1141	agctagtttgcacccctgctaaatcaagcactcttctgactcttctggagctaaattaca	1200		
Db	1141	AGCTAGTTTGGACCCCTGCTAAATCAAGCACTTCTTTTTCGACTTCTGAGACTTAATTACA	1200		
QY	1201	gttttgacaaaaatgctgacacaaacctgaatggcagacaaaggcatttccaatagtgtg	1260		
Db	1201	GTTTGTGACAAAAATGCTGACACAAACCATGAATGGCGAGACAAAGGCATTTCCAATAGTTG	1260		
QY	1261	ggagaccttccagaagcatcaaaagtttgagttgggctctcttcttggaggaggtgtt	1320		
Db	1261	GGAGCCTCTTCAGAAAGCATCAAAGTTTGGATTTGGGCTCTCTTTTGTGTGAGGGAGTTT	1320		
QY	1321	ctaattctaaacctgtaaagcgagcaggaagatatatttaattcaaaaggtgcctaatacac	1380		
Db	1321	CTAATTTCTAACTGTAAATGACGAGGAAGATATTTTAATTCAAAGGTGCCAATAACAC	1380		
QY	1381	ggacacaaatcaactccttgaattgatcacgacagagagaaacaaaggaggtggacagca	1440		
Db	1381	GGACCAAAATCAACTCCTTTTGAATTCATACGACAGAGAGAAAACAAGAGGAGGTGGACAGCA	1440		
QY	1441	aggaacactttcttatttgggaatctcaaccacatattcttcttggatagaatcagg	1500		
Db	1441	AGGAAACCTTTTCTTATTTTGGAATCTCAACCACATGATTCTTTGTTGTAGAGATGCAGG	1500		
QY	1501	ctcaaaaagtaatgcatgttcttcagcagaactgaattattcaactgccaatagcactcta	1560		
Db	1501	CTCAAAAAGTAAATGCATGTTTCTTCAGCAGAACTGAATTAATTCACCTGCCATATGACTCTA	1560		
QY	1561	aaacacaaatacgttaatgcctctaataagcaccatgactctagctctcttcttgggtat	1620		
Db	1561	AACACCAAAATACGTAATGCCTCTTAATGTAAGCACCACCTACTAGTGTCTTCTTGGTGTAT	1620		
QY	1621	attcttaatacactttagtggaaaaatcccttatttttcatcatgctgcctcaagtggtacca	1680		
Db	1621	ATTCTTACATACCTTTAGTGGAAAATCCCTTATTTCATCATGGCCTCCAAGTGGTACCA	1680		
QY	1681	gttctaagatgtctcttatttacctgagaagcaagatggaactgttttccctcttctc	1740		
Db	1681	GTTCTAAGATGTCCTTTGATTACCTTGAGNAGCAAGATGGAACCTGTTTTCTTCTTCTC	1740		
QY	1741	tgttgcaacatcctctacatccctcttcttcttattacaattcacatagttcttcttacc	1800		
Db	1741	TGTTGCAACATCCTCTACATCCCTCTCTTCTTATTACAATTCACATAGTCTCTTTATCAC	1800		
QY	1801	tgaattctcaacaaatatttctcactattggaacagagtcagctgttactagcaactg	1860		
Db	1801	TGAATTCCTCAACCAATATTTCTCTACTATTGAACAGGAGTCAGCTGTACTAGCAACTG	1860		

```
Qy 1861 ctccaagatagatgataaattccccctccacttccctgtagcacacctaataatttta 1920
Db 1861 CTCCAAGATAGATGATAAATCCCCCTCCACTTCTCTGTAGGACACCTGAATCATTTA 1920
Qy 1921 ttgtggttgaggaagctggagaattctcaccaaattgttcccaaatccttattcctcagctg 1980
Db 1921 TTGTGGTTGAGGAAGCTGGAGAATTTCTCACCAAATGTTCCCAAAATCCTTTATCCTCAGCTG 1980
Qy 1981 tgaaggtaaaattggaacatcactggaattggggtggaacatctgaacccaagaatttg 2040
Db 1981 TGAAGGTAAAATTGGAACATCACTGGGAATGGGTGGGAACATCTGAACCAAGAAATTTG 2040
Qy 2041 atgactctgtgatacttagaccaagcagagtgtaaaactcgaagtccttaaatcagcta 2100
Db 2041 ATGACTCTGTGATCTTAGACCAAGCAAGACAGTGTAAACTCCGAAGTCTTAATCAGGTA 2100
Qy 2101 aaaatttctctgtgcttagatgacatttagcccttaagattgggaagaatggttcgttaag 2160
Db 2101 AAAATTTCTCTTGGCTTTAGATGACATTTAGCCCTTAAGATTGGAAGAATGTTCTGTTAAG 2160
Qy 2161 tttagagtaattcacttcaggaagttaattggttcccaataatagcttccagattcattg 2220
Db 2161 TTTAGATTAATTCATCTCAGGAAGTTACTTGGTTCCCAATAATAGCTTCCAGTATTCAATTG 2220
Qy 2221 atttattctgcttccagactagaaatttctgaagagtcagtggggaagctagggc 2280
Db 2221 ATTTATTTCTGGCTTCCCAACTAGAAATTTTGTAAAGATCATGGGGGAAGCTAGGGC 2280
Qy 2281 taaccagaaaaataaaataatgggataaaaaatcggaactactgttttcccccta 2340
Db 2281 TAACCAGAAAAATAAATAAATGATGGATAAAAAATCGGAACACTACTGTTTCCCCCTA 2340
Qy 2341 gtcggagcacatcggg 2356
Db 2341 GTCGGAGCACATCCGG 2356

RESULT 2
AF001847 AF001847 2356 bp mRNA linear PRI 01-OCT-2001
LOCUS Homo sapiens lymphoid phosphatase Lyp2 mRNA, complete cds.
DEFINITION AF001847
ACCESSION AF001847
VERSION AF001847.1 GI:4100633
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2356)
Cohen, S., Dadi, H., Shaoul, E., Sharfe, N. and Roifman, C.M.
Cloning and characterization of a lymphoid-specific, inducible
human protein tyrosine phosphatase, Lyp
Blood 93 (6), 2013-2024 (1999)
99168989
PUBMED 10068674
REFERENCE 2 (bases 1 to 2356)
Roifman, C.M.
Direct Submission
AUTHORS Submitted (30-APR-1997) Immunology and Allergy, The Hospital For
Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,
Canada
FEATURES
source Location/Qualifiers
1..2356
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2120
/note="splice variant of lymphoid phosphatase Lyp1"
/codon_start=1
/product="lymphoid phosphatase Lyp2"
/protein_id="AAD00905.1"
/db_xref="GI:4100634"
/translation="MDQREILQKFLDEAQSQKITKEEFANFELKLKRSQTKYKADKTY
```

```
PTVAENAKNKKRKYDILPYDYSRVELSLITSDSESSYINANFKGVYGPXAYIAT
QGPISTLLDFWIMWEYSVLIIVMACMEYEMGKKCKERYWAEPGEQMLEFGEPSVSC
EAKRKSDDYIIRLKYKFNSETRIYQFHYKNPDDHVPSSIDFSLIREMRQPSLVGT
QOYELVYNVLELFKQMDVIIRDKHSGETSQAKHCLPEKNHTLQADSDSPNLPKSTT
KAAMNQORTKMEIKESSEDFRTSISAKIELVLHPAKSSTSFDFLELNYSPDKNA
DTYMKWOTKAPPIVGEPLQKHLJDLGSLLEFEGCSNKPVNAAGRFPNSKVTITRTKS
TPPELIOQRETKEVDKENFSLYESOPHDSCEVEMOAKQKVHVSSEALNTSLPYDSKH
OIRNASNVKHHSDSALGVYSYIPLVENPYESSMPPSCTSSKMSLDLPEKODGTVFPPSS
LLPTSSLSLFSYNSHDSLSLNSPTNLSLNGESAVLAPATRIDDDIIPPLPVRTPE
SFIVVEEAGFSPNPKVSLSSAVKVIKGTISLEWGGTSEPKKFDSDSVILRPSKSVKLRS
PKSGKNFSWL"
BASE COUNT 750 a 466 c 466 g 674 t
ORIGIN
Query Match 99.9%; Score 2352.8; DB 9; Length 2356;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 tccctcaacctacttatagactattttcttgcctgcagcatggaccacaagagaatttc 60
Db 1 TCCCTCAACCTACTATAGACTATTTTCTGCTCTGCAGCATGGACCAAGAGAAATTC 60
Qy 61 tgcagaagttcctgtagagcccaagcaagaaattactaaagagaggtttgcctaag 120
Db 61 TGCAGAAGTTCTCTGATGAGGCCCAAGCAAGAAAATTTACTTAAGAGGAGGTTTGCCTAATG 120
Qy 121 aattctgaagctgaaaaaggcaattctaccgaagtcaaggcagacaaacctatcctacaa 180
Db 121 AATTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAAGGCAGACAAAACCTATCTCTACAA 180
Qy 181 ctgtggtgagaatgccagaataatcaagaataatcaagaataatcaagaataatggccctatg 240
Db 181 CTGTGGCTGAGAATGCCAAGAAATATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATG 240
Qy 241 attatagccggtagaactatccctgataacccctctgagagattccagctacatcaatg 300
Db 241 ATTATAGCCGGTAGAATCTCCCTGATACCTCTGTATGAGGATTCAGCTACATCAATG 300
Qy 301 ccaacttcattaaaggagtttatggaccgaaggcttatattgcacccagggctcttat 360
Db 301 CCAACTTCATTAAAGGAGTTTATGGACCCCAAGGCTTATATTTGCCACCCAGGGTCTCTTAT 360
Qy 361 ctacaacctctctgactcttgaggagatttgggaataatagttccttcatcattgta 420
Db 361 CTACAACCTCTCTGACTCTCTGGAGGATGATTGGGAATATAGTGTCTCTTATCATTTGTTA 420
Qy 421 tggcatgcatgagtgatgaaatgggaagaaagtgtagcgctactgggctgagccag 480
Db 421 TGGCATGCTGAGTATGAATGGGAAGAAAGTGTGACGCTACTGGCTGAGCCAG 480
Qy 481 gagagatgcagctggaatttggcccttctctctgatactctgtagagctgaaaaaggaat 540
Db 481 GAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGAAAAAGGAAT 540
Qy 541 ctgattataatcaggactctaaaagttaagtcaatagtagaactcgaactatctacc 600
Db 541 CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTAGAAACTCGAACTATCTATCC 600
Qy 601 agtttcattacaagaattggccagaccatgatgtacacctctctatagaccctattcttg 660
Db 601 AGTTTCATTACAAGAAATTTGGCCAGACCATGATGTACCTTCATCTATAGACCTTATCTTCTG 660
Qy 661 agtcatctgggagtacgttggttaccagaaggatgacagctgttcccatatgcattcaact 720
Db 661 AGCTCATCTGGGATGTCAGTGTGTTTACCAAGAGATGACAGTGTGTCCCATATATGCAATCACT 720
Qy 721 gcagtgcgtgctggaaggaactggttatttgccttctgcttgcattgattatcattgattg 780
Db 721 GCAGTGCCTGCTGTGGAAGGACTGGTGTATTGTGCTGCTATTGTTGATTATACATGGAATGT 780
Qy 781 tgcataaagatgggataattcctcagaacttcagtggttttcagtttgatccgggaaatgc 840
```

Db 781 TGTCTAAAGATGGGATAATTCCTGAGAACTTCAGTCTTTTCAGTTTCATCCGGGAATGC 840
QY 841 ggcacagagccttcattagttcaaacacagagaacaatatgaactgggtctacaatgcctg 900
Db 841 GGACACAGAGCCCTTCATTAGTTTCAACCGCAGGAACAATATGAACATGGTCTCAATAGCTG 900
QY 901 tattagaactatttaagagacagatggatgttatcagagataaacaattctcggacacagaga 960
Db 901 TATTAGAACTATTTAAGAGACAGATGGATGTTATCAGAGATAAACATTCCTGGAAACAGAGA 960
QY 961 gtcagcaaaagcattgtattccttgagaaaaatcacactctccaagcagactcttattctc 1020
Db 961 GTCAAGCAAAAGCATTTGATTCTCTGAGAAAATCACACTCTCCAAGCAGACTCTTATTCTC 1020
QY 1021 ctaattaccaaaaagtaccacaaaagcagcaaaaaatgatgaaccaaaaagacaaaaa 1080
Db 1021 CTAATTTACCAAAAGTAGCCACAAAAGCAGCAAAAATGATGAACCAACAAAGACAAAAA 1080
QY 1081 tggaaatcaagaactctctctcttgacttttagctttctgaaataagtgcacaaagaag 1140
Db 1081 TCGAAATCAAGAATCTCTCTCTTTGACTTTTAGGACTTCTGAAATAGGTGCAAAAGAG 1140
QY 1141 agctagtttgcacctgctaataagcaactctctttgactttctggagctaaataaca 1200
Db 1141 AGCTAGTTTTCACCTGCTTAAATCAAGCACTTCTTTTGACTTTCTGGAGCTAAATTACA 1200
QY 1201 gttttgacaaaaatgctgcacacacacatgaatggcagacaaagcattctccaatagt 1260
Db 1201 GTTTTGCAAAAATGCTGCACAAACCATGAAATGGCAGACAAAGCATTTCCCAATAGTTG 1260
QY 1261 gggagcctcttcagaagcatcaaagtttgattggctctctctttgttgaggagatt 1320
Db 1261 GGGAGCCTCTTCAGAGCATCAAAGTTTGGATTTGGCTCTCTCTTTGTTGAGGGATGTT 1320
QY 1321 ctaattctaaacctgtaaatgcagcggaagagataatttttaattcaaaagggcgcaataaac 1380
Db 1321 CTAATTCTAAACCTGTAAATGCGACGAGAGAGATATTTTAAATCAAAGGTGCCAATAACAC 1380
QY 1381 ggacaaaatcaactctctttgaattgatcacagacagagaacaaagaggtggacagca 1440
Db 1381 GGACCAAAATCACTCTTTTGAATTTATACAGCAGACAGAGAAACCAAGAGGTGCACAGCA 1440
QY 1441 agaaaaactttctatttgggaatctcaaccacatgattcttctttgttagagatgcagg 1500
Db 1441 AGGAAAACCTTTCTTATTGGAATCTCAACCATCATGATTCTTCTTTGTAGAGATGCAGG 1500
QY 1501 ctcaaaaagtaatgcattgtttctcagcagaactgaattattcaactgcacatgactcta 1560
Db 1501 CTCAAAAAGTAATGCGATGTTTCTTCAGCAGAACTGAATATTCACTGCCATATGACTCTA 1560
QY 1561 aacacaaaatcacgtaaatgcctctaagttaagcaccatgactctagtgcctcttggtgat 1620
Db 1561 AACACCAAAATACGTAATGCGCTCTAATGTAAGCACCATGACTCTAGTGCCTCTGGGTAT 1620
QY 1621 attctacacactcttagtggaaaaatcctctatttttcatcatgctgcctccaagtggtaacca 1680
Db 1621 ATCTTTACATACCTTTAGTGGAAAATCCTTATTTTTCATCATGSCCTCCCAAGGTGACCA 1680
QY 1681 gttctaagatgctctctgatttaccctgagaagaagaatggaactgttttccctctctc 1740
Db 1681 GTTCTAAGATGCTCTCTGATTGTTTACCTTCAGAGCAAGTAAGTAAGTCTTCTCTCTCTC 1740
QY 1741 tgttgcaacactctctacatcctctctctctattacaaattcacatgattctttatcac 1800
Db 1741 TGTGCGCAACATCTCTACATCCCTCTTCTCTATTATCAAAATTCACATGATCTTTATCAC 1800
QY 1801 tgaattctcaacaaaatatttctcactattgaacaggagtagctgtactagcaactg 1860
Db 1801 TGAATTTCTCAACCAATATTTCCTCACTATTGAACAGGAGTCAGCTGTACTAGCAACTG 1860
QY 1861 ctccaagatagatgatgaaatccccctccactctctgtcacggaacactgaaatcattta 1920
|||||

Db 1861 CTCCAAGAGTAGATGATGATGAATCCCCCTCCACTCTCTGTACGGACACCTGAAATCATTTA 1920
QY 1921 ttgtggttaggaagctggagaattctcaccaaattctcccaaatcttaccatcagctg 1980
Db 1921 TTGTGTTTGGAGAGCTGGAGAAATTCACCAAAATGTTCCCAAAATCTTATCTCTCAGCTG 1980
QY 1981 tgaagtaaaaaattggaacatcacactggaatgggttggaacatctgaaccaaagaaatttg 2040
Db 1981 TGAAGCTAAAAATTTGGAACATCACCTGGAATGGGTGGAAACATCTGAACCAAAAGAAATTTG 2040
QY 2041 atgactctgtgatacttagacaagaagaagtgtaaaaaactccgaagctcctaaatcaggta 2100
Db 2041 ATGACTCTGTGATACTTAGACCAAGCAAGAGTGTAAAACTCCGAAGTCTCTAAATCAGGTA 2100
QY 2101 aaaaattctctggcttttagatgacatttagccctaaagatttgaaagaatgggttcgttaag 2160
Db 2101 AAAATTTCTTGGCTTTTAGATGACATTTAGCCCTTAGATTGGGAAGAAATGGTTCGTTAAG 2160
QY 2161 tttagagtaattcaacttcagggaattacttggttcccaataatagcttccagttattcattg 2220
Db 2161 TTTTAGAGTAATTCACCTTCAGGAAGTTTACTTGGTTCCCATAAATAGCTTCCAGTATTCTATTG 2220
QY 2221 atttattctggcttcccgagactagaaaattttgtaaaaagatcatgggggaagctagggc 2280
Db 2221 ATTTATTCTGGCTTTCCCGAGACTAGAAAATTTTGTAAAGAGTCATGGGGGAAGCTAGGGC 2280
QY 2281 taaccagaaaaataaaaaataatgggataaaaaatcggaactactgtttttccccccta 2340
Db 2281 TAACCAAGAAATATAAATAAATAATGGGATAAANAATCGGAACACTACTGTTTTCCCCCTA 2340
QY 2341 gtccggagcaatccgg 2356
Db 2341 GTCGGAGCACATCCGG 2356
|||||

RESULT 3

AX020261
LOCUS AX020261 3058 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO9936548.
ACCESSION AX020261
VERSION AX020261.1 GI:10044042
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3058)
AUTHORS Roifman,C.N.
TITLE Human lymphoid protein tyrosine phosphatases
JOURNAL Patent: WO 9936548-A 1 22-JUL-1999;
HSC RES DEV LP (CA); ROIFMAN CHAIM M (CA)

FEATURES

source
1..3058
/organism="Homo sapiens"
/db_xref="taxon:9606"
42..2468
/note="unnamed protein product"
/codon_start=1
/protein_id="GI:10044043"
/db_xref="GI:10044043"
/translation="MDQRIQLKFLDEAQSQKTKKEEFANEFKLKRQSTKYKADKTY
PTTVAENAKNIKKNRYKDIPLVDYGRVSLSLTSDSSYINANFIKGYGPKAYIAT
QGPLSTLLDFWRMIWEXSLIIVMACMEYEMGKKCYWAEPMQLEFPFVSVC
EAEKRSYIIRTLKVNSEFTTYQHYKNPDHVPSSIDPILLELWDVRCYQKH
DSVPICIHCSAGCGTGTCAIVDTMLLKDGIIPENFSVFLIREMTORPSLVQT
QROYELVNAVLELPKQMDVIRDKHSTESQAKHCIPKKNHTLOADSISNLPKSTT
KAAKMNQORTKMEIKRESSDFRTSEISAKHEELVHPKASTSDFLNLNTPDKNA
DTFMKQTKAFPIVGEPLQKHQSLDGLSLLFEGCSNKPVNAGRYFSKVPITTKS
TPFELIQRETKEDSKENFYSLEQPHDSCEVEMQAKVMHVSAAELNYSLEPDSKH
QIRNASNVKHHDSALGVSYIPLVENPYFSWPPSSKMSLDLPKQDQTVRPS
LLPTSTSLFSYNSHSSLSLNPTNISLLNOESAVLATAPRIDELIPPLVPRTPE
STIVVEEAGEFSPNVPKSLSSAVKVIKGTSLWEGGTSEPKFEDDSVILRPSKVLRS
PKSELHQDRSPPPPLPPLPFTLESFFLADEDCMQAQSIIETYSISYPTDMENSSKQTL

CDS

KTPGKSFTRSKSLILRNMKKKSICSNPPNKPAPESVQSNSSSFLNFGFANRFPKPKG
PRNPPTWNI*

BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

Query Match 89.0%; Score 2097; DB 6; Length 3058;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tccctcaactactatagactattttctgtctgcagcatggaccacaagagaattc 60
Db 1 TCCTCAACTACTATAGACTATTTTCTGCTGTCAGCATGGACCAAGAGAAATC 60
Qy 61 tgcagaagttccctggatggcccaagcaagaaattactaaagagaggttgcgaatg 120
Db 61 TGCAGAGTTCTCGGATGAGGCCCAAGCAAGAAANATTACTAAAGAGGATTTGCCAATG 120
Qy 121 aattctgaagctgaaaaggcaatctaccaagtacaaaggcagacaaacacctatcctacaa 180
Db 121 AATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGGCAGACAAAACCTATCCTACAA 180
Qy 181 ctgtgctgagaatgccaaagaatacaagaagaacagataaaggatattttgccctatg 240
Db 181 CTGTGGCTGAGAATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGCCCTATG 240
Qy 241 attatagccgggtagaaactccctgataacccctctgatgaggttcagctacatcaatg 300
Db 241 ATTATAGCCGGGTAGAACTATCCTGNTAACTCTGTATGAGGATTCGAGCTACATCAATG 300
Qy 301 ccaacttcattaaggaggtttatggaccacaaggcttatattgccaccacagggtcccttat 360
Db 301 CCAACTTCATTAAAGGGAGTTTATGGACCCCAAGGCTTATATGGCCACCAGGGTCTTTAT 360
Qy 361 ctacaaacctccctggactctgagagatgatttgggaataatagtctcttatcatgtta 420
Db 361 CTACAAACCCCTCGGACTCTCGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTA 420
Qy 421 tggcatgcatggagatgaaattgggaagaaaagtgtagcgtactggcgtgagccag 480
Db 421 TGGCATGCTGGAGTATGAATGGGAAGAAAAGTGTGAGCGCTACTGGGCTGAGCCAG 480
Qy 481 gagagatgcagctggaaatttggcccttctctgtatctctgtgaagctgaaaaagaaat 540
Db 481 GAGAGATGCAGCTGGAAATTTGGCCCTTCTCTGTATCTGTGAAGCTGAAAAAGGAAT 540
Qy 541 ctgattataatacaggactctaaaagttaagtccaatagtgaaactcgaactatctacc 600
Db 541 CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACACTCGAACTATCTACC 600
Qy 601 agtttcattacaagaattggccagacatgatgtacctctcatatagaccctattcttg 660
Db 601 AGTTTCATPACAGAAATGGCCAGACATGATGTACCTTCTATAGACCCCTATTCTTG 660
Qy 661 agctcatctgggttacgttgttaccagaagatgacagtggttcccatatgcatcaact 720
Db 661 AGCTCATCTGGGATGTACGTTGTACCAAGAGGATGACAGTGTCCCATATGATTCAC 720
Qy 721 gcagtgctggctggaagacttgggttatttggctatttggattatcacatgagatg 780
Db 721 GCAGTGTGCTGTGGAAGACTGGGTATTATTTGTGCTATTGTTGATTATACATGGATGT 780
Qy 781 tgcataaagatggataattccctgaaactcagttcttccagtttgcggttgcggaatgc 840
Db 781 TGTCTAAAGATGGGATTAATTCCTGGAACCTCAGTGTGTTTTCAGTTTGTATGATCCGGGAATGC 840
Qy 841 ggacacagagcccttcattagtctcaacgcaggaaacaatatgaactgggtctacaatgctg 900
Db 841 GGACACAGAGCCCTCATTAGTTCAACCGCAGGAACAATATGAACTGGTCTACATGCTG 900
Qy 901 tattagaactatttaagagacagatggatgttatacagagataaacattcttggaacagaga 960
Db 901 TATTAGAACATATTAAAGAGACAGATGGATGTTTATCAGAGATAAACATCTCTGGAACAGAGA 960

Qy 961 gtcaagcaagcattgtattctctgagaaaaatcacactctcccaagcagactcttattctc 1020
Db 961 GTCACGCAAGCAATGTATTCTTGAGAAAATCACACTCTCCAAGCAGACTCTATTCTC 1020
Qy 1021 ctaattaccaaaaagtcaccacaaaagcagcaaaaatgatgaaccaacaagaagacaaaa 1080
Db 1021 CTAATTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAGACAAAAA 1080
Qy 1081 tgaatacaagaatcttctctctttagacttttaggacttctgaaataagtgcaaaagaag 1140
Db 1081 TGAATAATCAAGAATCTTCTTCTTACCTTAGGACTTCTGAAATAAAGTGCAAAAGAAG 1140
Qy 1141 agtagtttgcaccctgctaaatcaagcactcttcttgacttcttgagctctggaactaaatca 1200
Db 1141 AGCTAGTTTGCACCTGCTAAATCAAGCACTTCTTTTGTGACTTCTGGAGCTAAATTA 1200
Qy 1201 gtttgacaaaaatgctgacacaaaccatgaaatggcagacaaaaggcatttcccatagtg 1260
Db 1201 GTTTTGACAAAATGCTGACACAACCATGAATGGCAGACAAAAGGCATTTCCAATAGTTG 1260
Qy 1261 gggagcctcttcagaagcatcaaaagttggatttgggctctcttcttcttggagggatgtt 1320
Db 1261 GGGAGCCTCTTCAGAAAGCATCAAAGTTGGATTGGGCTCTCTTTTGTGAGGGATGTT 1320
Qy 1321 ctaattcctaaacctgtaaatgcagcaggaagatatitttaattcaaaaggtgccataaac 1380
Db 1321 CTAAATCTAAACCTGTAAATGCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATTA 1380
Qy 1381 ggacaaaatcaactccttttgaattgacagcagagagaaaagggaggtggacagca 1440
Db 1381 GGACAAAATCAACTCCTTTTGAATTTGATACAGCAGAGAGAAAACCAAGAGGTGGAC 1440
Qy 1441 agaaaaacttttcttatttgggaatctcaaccacatgattcttcttcttggtagagatgc 1500
Db 1441 AGAAAACCTTTCTTATTGGAAATCTCAACCACATGATTTCTTTTGTAGATGCAGG 1500
Qy 1501 ctcaaaaagcaatgcatgttcttcagcagaactgaattattcactgccaatatgactcta 1560
Db 1501 CTCAAAAGTAATGCATGTTCTTCAGCAGAACTGAATTAATTCACCTGCCATATGACTCT 1560
Qy 1561 aacacaaaatcacgtaactgcctctaaatgtaagcaccatgactctagtctcttgggtat 1620
Db 1561 AACACAAAATACGTAATGCCTCTAAATGAAGCACCACCTAGCTAGTGTCTTTGGTGTAT 1620
Qy 1621 attcttacatacctttagtgaaaaatccttatttttcaatcattggcctccaagtggta 1680
Db 1621 ATTCTTTACATPACCTTTAGTGGAAAATCCTTTATTTTTCATCATGCGCCTCCAAAGT 1680
Qy 1681 gttctaagatgctctgtgattacctgagaagcaagatggaactgttttccctctctc 1740
Db 1681 GTTCTAAGATGCTCTCTTGATTTACCTTGAGAAAGCAAGATGGAAGTGTTTTCTCTCTC 1740
Qy 1741 tgttgcaaacactctctacatccctctctcttattacaaattcacatagttcttcttccac 1800
Db 1741 TGTGTGCAACATCCTCTACATCCTCTCTCTTATTACAAATTCACATAGTTCTTTATCAC 1800
Qy 1801 tgaattctcaacccaataatttctctactatttgaaccagagtgatgtgtactagcaactg 1860
Db 1801 TGAATTTCTCAACCAATAATTTCTCTACTATTGAACAGGAGTCAAGTGTACTAGCAACTG 1860
Qy 1861 ctccaaggtatgatgataatccccctccactctctctctctctcactgtacggacacctgaatcatt 1920
Db 1861 CTCCAAGGATGATGATGATAATCCCCCTCCACTTCTCTGTACGGACACCTCGAATCAATTA 1920
Qy 1921 ttgtggttggagagctggagaattctccacaaatgttcccaaatcttctatccctcagctg 1980
Db 1921 TTGTGTGTTGAGGAAGCTGGGAATTTCTACCAAAATGTTCCCAAAATCCTTATFCTCAGCT 1980
Qy 1981 tgaaggtaaaaattggaacatcactggaattgggttggaaatctgaaacccaagaatttg 2040
Db 1981 TGAAGGTAAAAATTGGAACATCCTACTGGAATGGGGTGGAACATCTGAAACCAAGAAATTTG 2040

Qy 2041 atgactctgtatacttagacacaagagagtgtaaaactccgaagtcctctaaatcag 2097
|||||
Db 2041 ATGACTCTGTGATACTTAGACCAAGAGAGTGTAATAACTCCGAAGTCTCTAAATCAG 2097
|||||

RESULT 4
AF001846 3058 bp mRNA linear PRI 01-OCT-2001
LOCUS Homo sapiens lymphoid phosphatase Lyp1 mRNA, complete cds.
DEFINITION AF001846
ACCESSION AF001846
VERSION AF001846.1 GI:4100631
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3058)
AUTHORS Cohen,S., Dadi,H., Shaoul,E., Sharfe,N. and Roifman,C.M.
TITLE Cloning and characterization of a lymphoid-specific, inducible
human protein tyrosine phosphatase, Lyp
JOURNAL Blood 93 (6), 2013-2024 (1999)
MEDLINE 99168989
PUBMED 10068674
REFERENCE 2 (bases 1 to 3058)
AUTHORS Roifman,C.M.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1997) Immunology and Allergy, The Hospital For
Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8,
Canada

FEATURES
source Location/Qualifiers
1..3058
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="thymocytes"
42..2468
/codon_start=1
/product="lymphoid phosphatase Lyp1"
/protein_id="AAD00904.1"
/db_xref="GI:4100632"
/transl_table="MDOREILQKFLDEAOSKKINKSEFANEFLKRBQSYKXKADKY
PTTVAENAKNKKRKYDILPYDSRVSELSLITSDSESSYINANFIKGYGPKRAYIAT
QGPTSLDLDLFWIMWEYSVLIIYMACMEYEMGKKERYWAEPMQLEFGFVSVC
EAEKKSDYIIRTLKXFNSETRTYOYFHYKNPDHPVSSIDILILINDVRCYQED
DSVPICHSACCGRTGICAVIYTWMLKDGITIPENFSVILIRMRTPQPSLVOT
QOEVLYNAVLELFKQMDVIRDKHSGTESOAKHCIPERNHTLQADSYPNLPSKVT
KAAKMNQORTKBEIKSSPDFRTSEISAKELVLHPAKSSTSFDFLELNYSPDKNA
DTTWKQTKAPPIVEGELOKHOSLDLGLLPEGCSNKPYNAGRYENSKVPIPTKKS
TPFELIQRETKEVDSKENSFYLESPHQDSFVEMQAQVNHVSASLINTSLPIDSXK
QIRNASNVKHHDDLSALGYIYIPLVENPYFSSWPPSPGTSKMSLDLPEKQDGVFSS
LLPTSSLSFYNSHDSLSNPTNISLLNQBSAVLATAPRIDDEIPPLPVRTPE
SFIIVEEAGEFSPNPKSLSSAVKVKIGTSLWEGTSEPKKDDSDVILRPSKVKLRS
PKSELHODRSSPPPLPPTLESPFLADEDCMQASITETYSTYPTWENSTSSKQFL
KTPCKSFTRSKLILRNMKKSICNSCPPNKPESVQSNSSSFLNFGFANRFSKPKG
PRNPPPTWNI"
BASE COUNT 1015 a 592 c 553 g 898 t
ORIGIN

Query Match 88.9%; Score 2093.8; DB 9; Length 3058;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2095; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tccctcaacctactatagactattttcttctctgcagcatggaccacaaagagaattc 60
|||||
Db 1 TCCTCAACCTACTTATAGACTATTTTCTTCTCTGTGAGCATGGACCAAGAGAAATC 60
|||||
Qy 61 tgcagaagtctctgatgagcccaagcaagaaaaattactaaagagaggtttgccaatg 120
|||||
Db 61 TGCAGAAGTCTCTGGATGAGGCCCAAGCAAGAAAAATTACTAAAGAGGAGTTTGCCAATG 120
|||||
Qy 121 aatttctgaagctgaaggaacatctaccagatacaggcagacaaaacctatctacaa 180
|||||
Db 121 AATTCTTGAAGCTGAAAGGCAATCTACCAGATACAGGACAGACAAAACCTATCTACAA 180
|||||

Qy 181 ctgtggctgagaaatgccaaagaatatcaagaaaaacagataataagataattttgcctatg 240
|||||
Db 181 CTGTGGCTGGAATGCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGCCTATG 240
|||||
Qy 241 attatagccgggtagaactatccctgataaacctctgtagagattccagctacatcaatg 300
|||||
Db 241 ATTATAGCGGTGAGAACTATCCCTGATAAACCCTCTGATGAGGATTCCAGCTACATCAATG 300
|||||
Qy 301 ccaacttcattaaaggagtttatggaccacaaaggcttatattgccaccagaggtctttat 360
|||||
Db 301 CCAACTTCATTAAAGGAGTTTATGGACCCAAAGGCTTATATTGCCACCAGGCTCCTTTAT 360
|||||
Qy 361 ctacaacctctggactctctggagatgattgggaatatagtgctcttcatcttgta 420
|||||
Db 361 CTACAACCTCTCTGGACTCTCTGGAGGATGATTGGGAATATAGTCTCTTATCATTTGTA 420
|||||
Qy 421 tggcatgcatggagtataaaatgggaaagaaagtgtgagcgtactgggctgagccag 480
|||||
Db 421 TGGCATGCTGAGTATGAAATGGGAAAGAAAAAGTGTGAGCGCTACTTGGCTGAGCCAG 480
|||||
Qy 481 gagagatgcagctggaatttggcccttctctctgtatctgtgaagctgaaaaaggaat 540
|||||
Db 481 GAGAGATGCAGCTGGAAATTTGGCCCTTCTCTGTATCTCTGAACTGAAAAAGGAAAT 540
|||||
Qy 541 ctgattataatacaggactctaaaagttaaagttcaatagtgaaactcgaaactatctacc 600
|||||
Db 541 CTGATTATATAATCAGGACTCTAAAAGTTAAGTTCAATAGTGAAGCTCGAACTATCTACC 600
|||||
Qy 601 agttctatacaagaattggccagaccatgatgtaccttcatctatagacctatctctg 660
|||||
Db 601 AGTTTCAATTACAAGAAATGGCCAGACCATGATGTACCTTCTATCTATAGACCTTATCTTG 660
|||||
Qy 661 agctcatctggagttaactgtttaccagaagatgaagtggtcccatgcatctact 720
|||||
Db 661 AGCTCATCTGGGATGTACGTTGTTACCAAGAGGATGACAGTGTCCCATATGCAATCTACT 720
|||||
Qy 721 gcagtgctggctgtggaaggactggtgtattgtgctattgtgtattatcatagatgt 780
|||||
Db 721 GCAGTGCTGGCTGTGGAAGGACTGCTGTTATTGTGCTATTGTGATTATACATGGATGT 780
|||||
Qy 781 tgcataaagatgaggaataatctctgagaaactcagtgctttcagtttgatcgggaaatgc 840
|||||
Db 781 TGCTAAAAAGATGGGATAATTCCTGAGAACTTCAGTGTGTTTCAGTTTGTATCGGGAATGC 840
|||||
Qy 841 ggacacagagcctcattagttcaaacgcgaggaacaatatgaactggtctacaaatgctg 900
|||||
Db 841 GGACACAGAGCCTTCATTAGTTTCAACCGCAGGAACAATATGAAGTGGCTTACATGGCTG 900
|||||
Qy 901 tattagaactatttaagagacagatggatgttatcagagataaaacattctctggaacagaga 960
|||||
Db 901 TATTAGAATATTTAAGAGACAGATGGATGTTATCAGAGATAAACATTTCTTGGAACAGAGA 960
|||||
Qy 961 gtcaagcaaaagcattgtattctctgagaaaaatcaacactctccaagcagactcttcttc 1020
|||||
Db 961 GTCAAGCAAAAGCATTGTATTCTGTAGAAAAATTCACACTCTCCAAGCAGACTCTTTATTCTC 1020
|||||
Qy 1021 ctaatttaccaaaaagtagtaccacaaagcagcaaaaatgatgaacacaaagacacaaa 1080
|||||
Db 1021 CTAATTTTACCAAAAAGTACCACAAAGGAGCAAAAATGATGNACCACCAAGGACAAAAA 1080
|||||
Qy 1081 tggaaatcaaaagaatcttctctctcttcttcttcttcttcttcttcttcttcttcttct 1140
|||||
Db 1081 TGGAAATCAAAAGAATCTTCTCTCTTTGACTTTTAGGACTTCTGAAATAAGTGC AAAAGAAG 1140
|||||
Qy 1141 agctagttttgcaccct 1200
|||||
Db 1141 AGCTAGTTTTCACCCCTCTAAATCAAGCCTCTCTTTTGTGAGTCTTTTGTGAGTCTTAAATTA 1200
|||||
Qy 1201 gttttgcaaaaaatgctacacacacatgaaatggcagacaaagcatttccaatagttg 1260
|||||
Db 1201 GTTTTGACAAAAATGCTGACACACACATGAATGGCAGACAAAGGCATTTCCAAATAGTTG 1260
|||||

Qy 1261 gggagcctcttcagaagcatcaaaagtgttgattggggtcctctcttttggtaggagatgtt 1320
Db 1261 GGGAGCCTCTTCAGAAAGCATCAAAAGTTTGGATTGGGCTCTCTTTTGGAGGATGT 1320
Qy 1321 ctaattctaaacctgttaaatgcagcaggaagaatatatttaattcaaaaggcgcaataaacac 1380
Db 1321 CTAATTCFAAACCTGTAAATGCAGCAGGAAGATATTTAAATTCAAAGGTGCCAATAACAC 1380
Qy 1381 ggaccacaaactcctcttttgaattgatcacgagagagaaccaaaggaggtgacacga 1440
Db 1381 GGACCAAAATCAACTCTCTTTGAAATGTACAGCAGAGAGAACAACNAGAGGTGGACACGA 1440
Qy 1441 aggaacactttcttatttggaaactcaccacacatgattctgttttttagagatgcagg 1500
Db 1441 AGGAAACCTTTCTTATTTGGAACTCAACCCACATGATTCCTGTTTGTAGAGATGCAGG 1500
Qy 1501 ctcaaaaagtgaatgcatgtttcttcagcagaaactgaattatcactgcacatgactcta 1560
Db 1501 CTCAAAAAGTAAATGCATGTTCTTTCAGCAGAACTGAATTAATTCATGTCATATGACTCTA 1560
Qy 1561 aacccaaaatcgttaatgcctctaatgtaaagcaccatgactctagtgctctgtgtgat 1620
Db 1561 AACACCAATAGTAAGTCCCTCTAATGTAAAGCACCATGACTCTAGTGTCTCTTGGTGAT 1620
Qy 1621 attctacataccttttagtggaataccttatttttcoactgagtcctccaagtgtgtaacca 1680
Db 1621 ATTCTTACATACCTTTAGTGGAAATCCTTATTTTTCATCATGTCCTGCTCCAAAGTGTACCA 1680
Qy 1681 gttctaagatctctgtatttacctgagaagaagaatggaactgttttttctctctctc 1740
Db 1681 GTTCTAAGATGTCTTGTATTACCTGAGAAGCAAGATGGAACGTGTTTTCTCTTCTTCTC 1740
Qy 1741 ttttgccacactcctctacatccctctctctctcttattacaattccacatagttctttatcac 1800
Db 1741 TGTGTCACCAATCTCTACATCCCTCTCTCTTATTACAAATTCATGATTCATTATCAC 1800
Qy 1801 tgaattctccaaaccaatatttctcactatttgaaccaggagtcagctgtactagcaactg 1860
Db 1801 TGAATTCFCAACCAATATTTCTCTACATTTTGAACCAAGATGGAACGTGTTACTAGCAACTG 1860
Qy 1861 ctccaagatagatgataaatccccctccacttctctctctgtacgagacactgaatcattta 1920
Db 1861 CTCCAAGATAGATGAATGAAATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1920
Qy 1921 tttgtgttgagagctggagaattctcaccacaaatgttcccaaatccttctcctcagctg 1980
Db 1921 TTGTGTTGAGAGCTGGAGAAATCTCACCAAATGTTCCTCTCTCTCTCTCTCTCTCTCTCT 1980
Qy 1981 tgaagggtaaaaattggaacatcactggaatgggtggaacatctgaaacaaagaatttg 2040
Db 1981 TGAAGGTAAAAATTGGAACATCACTGGAAATGGGTGGGAACATCTGAACCAAGAAATTTG 2040
Qy 2041 atgactctgtatacttagacaaagcaagagtgtaaaactccgaagtcctaatacag 2097
Db 2041 ATGACTCTGTGATCTTAGACCAAGCAAGAGTGTAAACTCCGAAGTCTTAAATCAG 2097

RESULT 5
AR149916 AR149916 3580 bp DNA linear PAT 08-AUG-2001
LOCUS
DEFINITION Sequence 1 from patent US 6228641.
ACCESSION AR149916
VERSION AR149916.1 GI:151114507
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3580)
AUTHORS Jallal, B. and Plowman, G.D.
TITLE Diagnosis and treatment of PTP04 related disorders
JOURNAL Patent: US 6228641-A 1 08-MAY-2001;
FEATURES Location/Qualifiers
1. .3580 source

BASE COUNT 1213 a 666 c 618 g 1083 t
ORIGIN
Query Match 88.1%; Score 2074.6; DB 6; Length 3580;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2090; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
Qy 1 tccctcaaccttaactatagactatttttctgtctgcagcatggaccacaaagaaattc 60
Db 12 TCCCTCAACCTTACTATAGACTATTTTCTGCTCTGCAGCATGACCAAAAGAAATTC 71
Qy 61 tgcagaagtctctgagtgcgcccacaaagcaagaataattactaaagagaggtttgccaatg 120
Db 72 TGCAGAAAGTCTCTGGATGAGGCCCAAGCAAGAAATTTACTTAAAGAGGAGGTTGCCAATG 131
Qy 121 aattctgaagctgaaaaggcaatctaccaaagtacaaggcagacaaacctatctctacaa 180
Db 132 AATTTCTGAAGCTGAAAAGGCAATCTACCAAGTACAAGGCAGACAAACCTTATCTCAAA 191
Qy 181 ctgtggtcgtgagaatgccagaataatcaagaaaaacagatataaaggatatatttgcctatg 240
Db 192 CTGTGGCTGAGAACCCCAAGAAATATCAAGAAAAACAGATATAAGGATATTTTGGCCCTATG 251
Qy 241 attatagccgggtagaactatccctctgataacctctgtagaggtatccagctcacatcaatg 300
Db 252 ATTATAGCCGGGTAGACTATCCCTGATACCTCTGTATGAGGATTCACGCTACATCAATG 311
Qy 301 ccaacttcattaaaggaggttatggccccaggcttatatttgcacccagggtccctttat 360
Db 312 CCAACTTTCATTAAGGGAGTTTATGGACCCCAAGGCTTATATGGCCACCAGGCTCTTTAT 371
Qy 361 ctcaacccctcctgacctctcgtgaggtatttgggaataatagtgctctatctatcttta 420
Db 372 CTACAAACCTCCTGGACTCTCGGAGGATGATTTGGGAATATAGTGTCTCTTATCATTTGTTA 431
Qy 421 tggcatgcagtgcagtgcagaaagtggaagaaagtgtagcgcactctgggctgagccag 480
Db 432 TGGCATGCATGAGTATGAATGGAAAGAAAGTGTGAGCGCTACTGGCTGAGCCAG 491
Qy 481 gagagatgcagctggaatttggcccttctctgtatcctgtgaagctgaaaaaggaat 540
Db 492 GAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCCTGTGAAGCTGAAAAAGGAAT 551
Qy 541 ctgattataaatcagagactctaaaaagttaagtccaatagtgaaactcgaaactatctacc 600
Db 552 CTGATTATATAATCAGGACTCTAAAAAGTTAAGTTCAATAGTGAACACTCGAACTATCTACC 611
Qy 601 agtttcattacaagaattggccagaccatgatgtacctctcatctatagaccctattctg 660
Db 612 AGTTTTCATTACAAGAAATGGCCAGACCATGATGTACCTTCATCTATAGACCTTATTTCTTG 671
Qy 661 agtcatctgggagtacgtttgttaccagaaggtatgacagtggttcccatatgcattcact 720
Db 672 AGCTCATCTGGGATGTACGTTGTTTACCAAGAGGATGACAGTGTTCCTCCATATGCATTCAC 731
Qy 721 gcagtgcgtgcgtggaaggactggtttatttgcattgtattgttgattatatacatggatgt 780
Db 732 GCAGTGTGGCTGTGGAAGGACTGGTGTATTTGTGTGTA---TTGATATATACATGGGATGT 788
Qy 781 tgcataaagatgggataattcctcagaaactcagttcttcagtggttgcggaagaaatgc 840
Db 789 TGCTAAAGATGGGATTAATCTCTGAGAACTTCAGTGTGTTTTCAGTTCGCGGAATATGC 848
Qy 841 ggacacagagccttcattagttccaacgcaggagcaacaatatgaactgggtctacaatgctg 900
Db 849 GGACACAGAGGCCCTTCATTTAGTTCAACGCAGGAACAATAATGAACCTGGTCTTACAATGCTG 908
Qy 901 tattagaactatttaagagacagatggtgtttatcagagataaacattcttgggaacagaga 960
Db 909 TATTAGAACTATTTAAGAGACAGATGGATGTTTATCAGAGATAAACATCTTGGGACAGAGA 968

QY 961 gtaagcaagcattgtattcttgagaaaaatcacactctccaagcagactcttattctc 1020
Db 969 GTCAAGCAAGCATTGTATTCTGTGAGAAAAATCACACTCTCCAAGCAGACTCTTATTCTC 1028
QY 1021 ctaatttaccaaaagtaccacaaaaagcagcaaaaaatgatgaacccaagaagacaaaa 1080
Db 1029 CTAAATTTACCAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAGGACAAAAA 1088
QY 1081 tggaaaataaagaattcttctcttgacttttaggactctgaataaagtcaaaaagag 1140
Db 1089 TGGAAATCAAGAAATCTCTCTCTTTGACTTTTAGGACTTCTGAATAAGTGCAAAAGAG 1148
QY 1141 agtagttttgcacctgtataatcaagcactcttttgactttctggagctaaaattaca 1200
Db 1149 AGCTAGTTTGGACCTGCTAAATCAAGCACTTCTTTTGACTTTCTGGAGCTAAATTACA 1208
QY 1201 gttttgacaaaaatgctgacacaaacatgaataaggcagacaaaggcatttccaatagt 1260
Db 1209 GTTTTGACAAAAATGCTGACACAAACCATGAAATGGCAGACAAGGCATTTCCAATAGTTG 1268
QY 1261 gggagcctctcagaagcatcaaaagtgttgatttggctctcttcttcttggaggaatgtt 1320
Db 1269 GGGAGCCTCTTCAGAGCATCAAAAGTTTGATTGGGCTCTCTTTTGTGGGGATGTT 1328
QY 1321 ctaattctaaacctgtataatgcagcaggaagatatttttaattcaaaaggtgccaaatacac 1380
Db 1329 CTAATTTCAAACCTGTAATGCAGCAGGAAGATATTTTAATTCAAAGGTGCCAATAACAC 1388
QY 1381 ggacaaaataaacctcttttgattgatacagcagagagaaacaaaggaggtggacagca 1440
Db 1389 GGACCAAAATCAACTCTCTTTTGAATTGATACAGCAGAGAGAAACCAAGAGGTGGACAGCA 1448
QY 1441 aggaataactttctatttggaaatcaccacatgatcttctgtttttagagatgcag 1500
Db 1449 AGGAAACTTTCTTATTGGAAATCTCAACCATGATCTCTGTTTGTAGAGATGCAGG 1508
QY 1501 ctcaaaagaatgcattgtttcttcagcagaactgaattattcaactgccatatgactcta 1560
Db 1509 CTCAAAAGTAATGTCATGTTTCTTCAGCAACTGAATATTTCACGTGCCATATGACTCTA 1568
QY 1561 aacacaaaatagtaagctctaatgtaaagcaccatgactctagctcttctgtgtat 1620
Db 1569 AACACCAATACGTAATGCTCTAATGTAAAGACCATGACTCTAGTGTCTCTTGGTGTAT 1628
QY 1621 attctacatacctttagtggaaaatccttatttttcatatgcttccatcgaagtggtacca 1680
Db 1629 ATTTCTACATACCTTTAGTGGAAATCCTTATTTTTCATCATGCGCTCCAGTGTACCA 1688
QY 1681 gttctaagatgtctcttgattttacctgagaagcaagatggaactgttttctctctctc 1740
Db 1689 GTTCTAAGATGTCTCTTGATTTACCTGAGAAGCAAGATGGAAGTGTTTTCCCTTCTCTC 1748
QY 1741 tgttgcaaacatcctctacatccctctctcttctatttacaattcaatagttctttttatcac 1800
Db 1749 TGTGTGCAACATCCTCTACATCCCTCTCTCTTATTACAAATTCACATGATCTTTTATCAC 1808
QY 1801 tgaattctcaaccaaataattctctcactattgaaccagagtgctgtactagcaactg 1860
Db 1809 TGAATTTCCACCAATATTTTCTCTACTATTGAACAGGAGTCAGCTGTACTAGCAACTG 1868
QY 1861 ctccaaggtagatgatgaaatccccctccactctctctgaagacacactgaaatcatatta 1920
Db 1869 CTCCAAGGATAGATGATGAAATCCCCCTCCACTTCTCTGTACGACACACCTGAAATCATTTA 1928
QY 1921 ttgtggtgaggagctggagaattctcaccaaaatgttccccaaatccttatctcagctg 1980
Db 1929 TTTGTGTTGAGGAAGTGGAGAAATTCACCAAAATGTTCCCAAAATCCTTATTCCTCAGCTG 1988
QY 1981 tgaaggtaaaaattgaaacatcatcgaatggggtggaactctgaacacaaagaaatttg 2040
Db 1989 TGAAGGTAAAAATTTGAAACATCACTGGAATGGGGTGGACATCTGAACCAAGAAATTTG 2048
QY 2041 atgactctgtacttagaccagaagcagaagtgtataaactccgaagctcctaaatcacg 2097

Db 2049 ATGACTCTGTGATCTAGACCAAGCAAGAGTGTAACAACTCCGAAGTCTCTAAATCAG 2105
RESULT 6
AF150732 2877 bp mRNA linear PRI 04-MAY-2000
LOCUS Homo sapiens protein tyrosine phosphatase (PTP) mRNA, complete cds.
DEFINITION AF150732
ACCESSION AF150732
VERSION AF150732.1 GI:7688662
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2877)
AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2877)
AUTHORS Huang,C., Zhang,C., Wu,T., Peng,Y., Gu,Y., Zhang,L., Jiang,C.,
Li,Y., Han,Z., Wang,Y., Chen,Z. and Fu,G.
DIRECT SUBMISSION
TITLE Submitted (12-MAY-1999) Chinese National Human Genome Center at
JOURNAL Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Location/Qualifiers
source 1..2877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
gene 1..2877
/gene="PTP"
CDS 71..2470
/gene="Ptp"
/codon_start=1
/product="protein tyrosine phosphatase"
/protein_id="AAF67472.1"
/db_xref="GI:7688663"
/translation="MDQREILQKFLDEAQSKKITKEEFANEFLEKLKROSTKYKADKY
PTTVAEPKNINKNRYKDIPIYDSRVELSLTSDSSYINANIKGYGPEAIAT
QGPJSTLLDFWRMTWEYSVLINMACMEYEMGKRCRYAEAPGEMOLEFGPFVSVC
EAEKRKSDYIIRTLKFNSETRTIYQHYKNWPDHVPSSIDPLEIWDVRCQTQED
DSVPTICHCAGCGRTGICADYTWMLIKDGIIPENFSVFLIREMTQRSLVQTQ
ROYELVYNVLELFRKMDVIRDKHSGTSEAKHCIPKKNHTLOADSYSPNLPKSTTK
RAKMNQORTKWEIKESSDFERTSEISAKEDVSLAPENQAFECISEVNSSPAKMLPH
HEMPAKAFQIVGSSSEASKLNASLCLGNSNLVCSKFINPCROYDQINSFELIQ
QENQGDGSKTFLFESQPHDSCFVEMQAKQVNHVSSALNYSLPYDSKHQIRNASNVK
HDDSGALGVSYIPLVENPYSWPPSGTSSKMSLDLPEKODGYFPSPSLTSTSL
FSYYNHSDLSLNSPTNISRSRTEQESAVLATAPRIDDEIPLHFLGTPESFTVVEAG
FESPVPKSLSSAVKVKICTGLEWGTSEPKKFDSDVILRPSKSVKLRSKSELHQDR
SSPPPLPERLESIFLADEDCMQAQSIETYSTYPTDMENSTSSKQTLKTPGKSFTR
SKSLKILRNMKKSCINCSPPNPAESVSNSSSFLNFGFANRFSPKPRPPPTWN
I"

BASE COUNT 944 a 583 c 523 g 827 t
ORIGIN

Query Match 74.7%; Score 1761; DB 9; Length 2877;
Best Local Similarity 94.0%; Pred. No. 0;
Matches 1978; Conservative 0; Mismatches 100; Indels 27; Gaps 13;

QY 1 tccctcaactctatagactattttcttctctgcagctggaccacaaagaatic 60
Db 30 TCCTCAACTACTTATAGACTATTTTCTGCTCTGCAGCATGGACCAAGAGAATTC 89
QY 61 tgcagaagttctctgtagagggcccaagcaagaaaattactaaagagaggtttgc 120
Db 90 TGCAGAAGTTCTCTGGATGAGGCCCAAGCAAGAAAATTACTAAAGAGGAGTTTGC 149
QY 121 aattcttgagctgaaagggcaatctcacaagtcagaagcagacaaacctctcata 180

Db 150 AATTTCTGAAGCTGAAAGGCAATCTACCAAGTACAAAGCAGACAAAACCTATCTCAAA 209
Qy 181 ctgtgctgagaaatgcccaagaatatcaagaaaaacagagataaagatatatttgcctatg 240
Db 210 CTGTGGCTTGAGAGCCCAAGAAATATCAGAAAAACAGATATAGGATATTTTGGCCATG 269
Qy 241 attatagccgggtagaactatccctgataaactctgatgaggtattccagctacatcaatg 300
Db 270 ATTATAGCCGGGTAGAACTATCCCTGTATACCTCTGTATGAGGATGCCAGCTACATCAATG 329
Qy 301 ccaacttcattaaaggagatttataggaccccaaggcttatattgcacccagggctctttat 360
Db 330 CCAACTTCATTAAAGGAGTTTATGGACCCGAGGCTTATATTTGCCACCCAGGCTCCTTAT 389
Qy 361 ctacaacccctcctggactctggagatgatttgggaatatagtgctcttatctatgttta 420
Db 390 CTACAAACCTCCTGGACTTCTGGAGGATGATTTGGGAATATAGTGTCTTATCATATGTTA 449
Qy 421 tggcatgcatggagatgaaatgggaagaaaaagtgtagcgtactctggcgtgagccag 480
Db 450 TGGCATGCTGAGATGATGAATGGAAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAG 509
Qy 481 gagagatcagctggaatttggcccttctctgtatcctgtgaagctgaaaaaggaat 540
Db 510 GAGAGATCAGCTGGAATTTGGCCCTTTCTCTGTATCTGTGAAGCTGAAAAAGGAAT 569
Qy 541 ctgattataatcaggactctaaaagttaaagttcaatagtagaaactcgaaactatctacc 600
Db 570 CTGATTATATATCAGGACTCTAAAGTTAAGTTCAATAGTGAACCTCGAACTATCTACC 629
Qy 601 agtttcattacaagaattggccagaccatgatgtaccctctcatctatagaccctattctg 660
Db 630 AGTTTCAATACAGAATTTGGCCAGACCATGATGTACCTTCATCTATAGACCCCTATCTTG 689
Qy 661 agctcatcggagatgaactgttaccagaagagatgacagtggttcccatatgcattcaact 720
Db 690 AGCTCATCTGGGATGTCGGTTGTACCAAGAGGATGACAGTGTTCCTCCATATGCAATCACT 749
Qy 721 gcaatgctgctggaagactgatttattgtctattgtctattgttattgattatcatggatg 780
Db 750 GCAGTGTCTGGCTGTGGAAGGACTGGTGTATTGTGCTA---TTGATTATACATGGATGT 806
Qy 781 tgcataaagatgggataattcctgagaacttcagtggttttcagtttgatccgagaaatgc 840
Db 807 TGCTAAAGATGGGATAATCTCGAAGCTTCAGTGTCTTTCAGTGTGATCGGGAAATGC 866
Qy 841 ggacagagagccttcattagttcaaacgcgaggaacaataagaactggtctacaatgctg 900
Db 867 GGACACAGAGGCTTCATTAGTTCAACGCAGGAACAATATGAACCTGGTCTACAAATGCTG 926
Qy 901 tattagaactattaagagacagatggatgttatcagagataaacattcttgaaacagaga 960
Db 927 TATTAGAACTATTAAAGACAGATGGATGTTATCAGAGATAAACAATTCGTGGAACAGAGA 986
Qy 961 gtcaagcaaacgattgtattcctgagaaaaatcacactctccaagcagactcttattctc 1020
Db 987 GTCAGCAAAAGCATGTATTCTTGAGAAAAATCACACTCTTCAGCAGACTCTTATCTC 1046
Qy 1021 ctaatttaccaaaagtaccacaagaagcagcaaaaatgatgtgaacccaagaagcaaaaa 1080
Db 1047 CTAATTTTACCAAAAAGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAAGGACAAAA 1106
Qy 1081 tggaaatcaagaactcttctccttgacttttagactcttgaaataagtgcaaaaag 1140
Db 1107 TGGAAATCAAGAATCTTCTCTTGTGACTTTTAGACTCTTGAATAAGTCAAAAGAAAG 1166
Qy 1141 agctagttttgacccctgctaaaacagcactctcttttttgactttcttgagcgtataattaca 1200
Db 1167 ACGTAAGTTTGGCCCGG---AAACACAGGCTTTTGTGCTTTCGGAAGTAATTTCCA 1223
Qy 1201 gttttgacaaaatgctgacacaacatgaaatggcagacaaagcatttccaatgattg 1260
Db 1224 GTTTTG---CAAAAATGTTGCCCCACCAATGAATGCCAG-CAAAGGCTTTCCAAATAGT-- 1278

Qy 1261 gggagcctcttcagaagcatcaaaagtttggatttgggctctcttctgttggaggtgtt 1320
Db 1279 GGGAAAGCTCTTCAGAAAGCATCAAG-TTGAATTTGGGCTCTCTTTGTGTT---AGGAATGT 1334
Qy 1321 ctaatttcaacctgtaaatgcagcaggaagatatatttcaaaagtgcaataaacac 1380
Db 1335 CTAATTTCTAACCTTGT--ATGCAGCAGAAAGATATTT--AATCCAAGGTGCCAAT-ACCC 1389
Qy 1381 ggacccaaatcaactcccttttgaattgatacagagagagagaaaaacaaaggagtgacacga 1440
Db 1390 GGACCAAAATCAACT-CTTTTGAATTTGATCCAGCA---AGAGAACAAGGAGGTGACAGCA 1445
Qy 1441 agaaaaacttttcttatttggaaatctcaaccacatgatcttctgttttgaagatgacag 1500
Db 1446 AGAAAACCTTCTTATTT---GAATCTCAACCAACATGATCTTGTGTTGTAGAGATGAGG 1502
Qy 1501 ctcaaaagtaatgcattgtttcttcagcagaaactgaattattctactgcatatgactcta 1560
Db 1503 CTCAAAAGATATGCATGTTCTTTCAGCAGAACTGAATTTATTCATCTGTCATATGACTCTA 1562
Qy 1561 aacacaaaatcgttaatgcctcttaattgaaagacccatgactctagtccttctgttat 1620
Db 1563 AACACCAAAATACGTAATGCCCTTAATGTAAAGCACCATGACTCTAGTGTCTCTTGGTGAT 1622
Qy 1621 attctacataccttttagtgaaaaatccttatttttcatcatggtccctcaagtgtgaccca 1680
Db 1623 ATTTCTTACATACCTTGTAGTGGAAATCTTATTTTTCATCATGCGCTCCCAAGTGGTACCA 1682
Qy 1681 gttctaagatgctcttctgtatttacctgagaagcaagatggaaactgttttctctctctc 1740
Db 1683 GTTCTAAGATCTCTCTTGTATTACCTGAGAAGCAAGATGGAACCTGTTTCTCTTCTCTC 1742
Qy 1741 tgttgcaacaactctctacatccctctctctctctctcttattacaattcacatagttctttacac 1800
Db 1743 TGTTGCCAACATCTCTACATCCCTCTCTCTTATTATTAATTCATGATGATCTTTATCAC 1802
Qy 1801 tgaattctcaaaccaatatttctcactattgaacagagagctgactgactagcaactg 1860
Db 1803 TGAATTTCCAAACCAATATTTTCTCAGTATTGAACAGGAGTCACTGCTACTAGCAACTG 1862
Qy 1861 ctcaagagatagatgataatccccctccactctctctgtacgacacactgaaatcattta 1920
Db 1863 CTCAAGGATAGATGATGAATCCCTCCACTTCTCTGTACGGGACACCTGAATCATTTA 1922
Qy 1921 ttgtggttgaggaagctggagaaattctcacaataatgttcccaaatcttctcctcagctg 1980
Db 1923 TTGTGGTTGAGGAAGCTGGAGATTTCCACCAAAATGTTCCCAAAATCTTATCTCCAGCTG 1982
Qy 1981 tgaaggtaaaaattggaacatcactggaatgggtggaacatctgaaccaaagaatttg 2040
Db 1983 TGAAGGTAAAAATTGGAAACATCACTGGAATGGGTGGAACATCTGAAACCAAGAAATTTG 2042
Qy 2041 atgactctgtactactagaccaaagcagagtgtaaaactccgagctcctaatacagga 2100
Db 2043 ATGACTCTGTGATCTACTAGACCAAGCAAGAGTGTAAAACTCCGAAGTCTTAATCAGAAC 2102
Qy 2101 aaat 2105
Db 2103 TACAT 2107

RESULT 7

MUSPROTYPH
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism

MUSPROTYPH 2734 bp mRNA linear ROD 27-APR-1993
Mouse protein tyrosine phosphatase (702pep) mRNA, complete cds.

M90388

protein-tyrosine phosphatase.

Mus musculus cDNA to mRNA.

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 2734)
Matthews, R.J., Bowne, D.B., Flores, E. and Thomas, M.L.
Characterization of hematopoietic intracellular protein tyrosine
phosphatases: Description of a phosphatase containing an SH2 domain
and another enriched in proline-, glutamic acid-, serine-, and
threonine-rich sequences
Mol. Cell. Biol. 12, 2396-2405 (1992)

FEATURES	source	Location/Qualifiers
		1. .2734
		/organism="Mus musculus"
		/db_xref="taxon:10090"
		1. .2734
gene		/gene="70zpep"
		113. .2521
		/gene="70zpep"
CDS		/EC_number="2.7.1.112"
		/codon_start=1
		/product="protein tyrosine phosphatase"
		/protein_id="AAA39994.1"
		/db_xref="GI:200523"
		/translation="MDQREITQLLKEAQKKLNSEFESEFLKKRQSTKYYKADKIY PTTVAQRKNIKNRYKDIPLYPDHSIVELSLTSDSDSYINASFTRGVYGPAYATY QGPLSTLTDWFEMIEWYRLTVTMACMEFEMGKKCYERWAPGPTQLQFGPFSTSC QAGKSDYKIRPLKAFKNETRIIYQFHYKNWPDHDVPSSIDPLIQLWDMRCYQED EAPCICHSGAGRTGVCICAVDYITWMLLKNDHPNFVSNLIGIQEMRTQPSLVTQ QYELYYSVLLEFKRHMVDISNHLGREITQAQCSPEQSLTVEADSCPLDLKPNAM DKVTTNQHQGAESGSSGLGRLSTMTNAEELVLSHSAKSPSFCNLCLMCCGNR KAVITTEQARASVPVGPLOKYSQSLDFGMLFGSCPALPINTADRYHNSKGPVKRT KSPTFELIQORNTDLNADGDFSCLEQHEHYSLRELOVRVAHVSSSEELNYSLEPCA CDASCYPRHSFGALRVHLIYTSLAEDYPHFSPPNSADSKMSFDLPKQDQATSPGALL PASSTSYFSPHDSPLWNLTISFSPPLNQETAVEAPRSTQDEIPLPPLEPTSPSF IVVEEAGEPSRPVTESLPLVTVFGASEPCEGSEMKSHDVSFTPSKNKYLKPKSDR HDSQSPPLPRLPTELSFTLADBDCAQAAVQTSYTSPTETTNSTSSKQTLRTPKGS FTFSKSLKIFRNMKKSCVQNSSPSKPTERYOPKNSSSFLNFGNGNRFPSKPGPNRNPES

BASE COUNT	787 a	629 c	601 g	717 t	
ORIGIN					
Query Match	53.5%;	Score 1261;	DB 10;	Length 2734;	
Best Local Similarity	77.3%;	Pred. No. 9e-280;			
Matches 1637;	Conservative 0;	Mismatches 435;	Indels 45;	Gaps	7;
Qy	1	tcctcaacctactatagacattttctgtctgcagatgcacaaagagaaattc	60		
Db	78	TCCCTTAACCTACTACAGAC-----TCITGTCTCGGCATGCACCAAGACAATTC	131		
Qy	61	tgcagaagtctctggatgagggcccaagcaagaaataactaaagagaggttgcccaatg	120		
Db	132	TGCAGCAACTACTGAAAGAAGCCAGAGAAAAGAAACTCAACAGTGAGGAGTTTGCCAGTG	191		
Qy	121	aattctgaagctgaagaagcaatctaccaagtacaaggcagacaaaacctatcctacaa	180		
Db	192	AATTCTTGAGCTGAAAGGCAATCCACCAAGTACAAGCGCGACAAAATCTATCCTACAA	251		
Qy	181	ctgtggctggaatgccaagaatatcaagaaaaacagatataaagatatatttgcctatg	240		
Db	252	CTGTGGCTCAGAGCCCAAGNAATATCAAGAAAAACAGATACACAGGATATTTTGCCCTATG	311		
Qy	241	attatagccgggtagaactatccctgataacctctgatggagattccagctacatcaatg	300		
Db	312	ATCACGCTGGTAGAGCTGTCCTGTAACTTCCGATGAGGATTCAGTTATATCAATG	371		
Qy	301	ccaacttcattaaagggagtttatggaccaaggcttatattggaccacagggctcctttat	360		
Db	372	CCAGCTTTATTAAAGGGTGCTATGACCCCAAGGCTTATATTGCTACCCAGGGCTCTTTAT	431		
Qy	361	ctacaacctctggacctctggaggtatattgggaatatatagttccttatcatttta	420		
Db	432	CTACAACCTCTCGACTCTTGAGGAGTAGATTGGGAGTACCGCATCTTGCTATTGTCA	491		
Qy	421	tggcatcgatggagtatgaaattgggaagaaaaaagttggacgcgtactggcgtgagccag	480		

Db 1569 GGGAGCTGCAGGTGCAAAAGAGTGGCCCATGTTCTTTCAGAAAGAGCTGAATTATTACATGCG 1628
Qy 1549 catatgactctaaacaccaaatacgtactcctaaatgttaagacaccatgaactctagt 1608
Db 1629 CTGGTGCTT-----GTGATGGCTGCTGTGTGTCGCCCGGCACAGCCCGCGG 1673
Qy 1609 ctcttggtgtattctacatcacctcttagtggaataacccctatttttcatcatcgccctc 1668
Db 1674 CTTTGAGAGTGTCATGTACACATCTTTAGCGGAAGATCCTTATTATTTCATCATCCCTC 1733
Qy 1669 caagtggtaccagttctcaagtgtctctgtatttaccttgagaagaagaatggaactgttt 1728
Db 1734 CCAATAGTCTGCTGATTCGAAGATGCTTTTGTATCTGCTTGAAGAACAGGATGGAGCACTT 1793
Qy 1729 ttccttctctgttgccaaacatctctacatccctctctcttatttattacaattcacata 1788
Db 1794 CCCCTGGGCTCTATTGGCAGCTCTTCTACACCTCTCTTTATYATGAAACACCCACAGC 1853
Qy 1789 gttctttatcactgaattctccaaacattttctcactatttgaaccaggagtcagctg 1848
Db 1854 ACTCCCTAGTATGAACACTCTGACCACTTTTCCCCACCCTTAAACCAAGAGACAGCTG 1913
Qy 1849 tactagaactgctccaaagtagatgatgaataccccccctccactctctctgtacaggacac 1908
Db 1914 TAGAAGCTCTTCTCGGAGGACAGATGATGAATCCCGCCGCCACTCCTCGAACGGACAC 1973
Qy 1909 ctgaatcattatgtgtgaggaagctggagaattctcaccataatgttcccaaatcct 1968
Db 1974 CCGAGTCTTTATGTGTGTGAGGAGCCGGAGAGCCCTCACCAGCTGTTACCGAATCCT 2033
Qy 1969 tatctcagctgtgaagtaaaattggaacatcactggaatgggtgggaacatctgaac 2028
Db 2034 T-----ACCTCTGTGTGTAACATTTGGAGCATCACCAGATCGATGGGACATCTGA-- 2085
Qy 2029 caaagaatttgatgactctgtgacttacttagaccagaagaagtgtaaaactcgaagctc 2088
Db 2086 -AATGAAGAGGCTGACTCTGTAGGTTTACACCAAGCAAGAATGTGAATCCGAAGTC 2144
Qy 2089 ctaaatcaggtaaaaat 2105
Db 2145 CCAATCAGATCGACAT 2161

RESULT 8
LOCUS AL137856/c
DEFINITION Human DNA sequence from clone RP5-107303 on chromosome 1p13.1-13.3,
complete sequence.
ACCESSION AL137856
VERSION AL137856.24 GI:14970654
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143618)
Glithero.R.
Direct Submission
Submitted (17-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On Jul 19, 2001 this sequence version replaced gi:14586004.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-107303 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

This sequence is the entire insert of clone RP5-107303 The true
right end of clone RP11-215C20 is at 30467 in this sequence. The
true right end of clone RP11-324J2 is at 119169 in this sequence.
Location/Qualifiers

FEATURES
source

1. 143618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p13.1-13.3"
/clones="RP5-107303"
/clone_lib="RPCI-5"
257..322
/note="MER5A repeat: matches 41..109 of consensus"
763..11053
/note="Alusx repeat: matches 1..312 of consensus"
1348..1698
/note="L2 repeat: matches 2070..2417 of consensus"
1919..1969
/note="L2 repeat: matches 2655..2705 of consensus"
2305..2446
/note="MIR repeat: matches 40..176 of consensus"
5652..5927
/note="L1ME repeat: matches 5524..5813 of consensus"
6295..6428
/note="BC200 repeat: matches 2..147 of consensus"
6653..6700
/note="8 copies 6 mer acacac 97% conserved"
7925..7980
/note="28 copies 2 mer aa 75% conserved"
8071..8369
/note="Alusx repeat: matches 1..296 of consensus"
9048..9342
/note="Alusx repeat: matches 1..298 of consensus"
9793..10559
/note="L1MC1 repeat: matches 5526..6320 of consensus"
10574..10730
/note="MIR repeat: matches 95..262 of consensus"
10759..10907
/note="MIR repeat: matches 2..139 of consensus"
10908..11219
/note="AluJo repeat: matches 1..296 of consensus"
11220..11315
/note="MIR repeat: matches 139..237 of consensus"
12210..12464
/note="Alusg repeat: matches 1..291 of consensus"
12916..13149
/note="MER58B repeat: matches 70..341 of consensus"
13150..13435
/note="Alusx repeat: matches 1..298 of consensus"
13458..13805
/note="L1MD2 repeat: matches 4851..5213 of consensus"
13806..14115
/note="AluJb repeat: matches 1..299 of consensus"
14118..14430
/note="Alusx repeat: matches 1..312 of consensus"
14331..14563
/note="L1MD2 repeat: matches 5213..5314 of consensus"
14565..14859

```
repeat_region /note="AluSg repeat: matches 1. .294 of consensus"
14861. .14888
repeat_region /note="14 copies 2 mer ta 92% conserved"
14910. .15215
repeat_region /note="L1MD2 repeat: matches 5302. .5609 of consensus"
15230. .15560
repeat_region /note="MER2 repeat: matches 2. .345 of consensus"
15561. .16146
repeat_region /note="L1MD2 repeat: matches 5643. .6210 of consensus"
16270. .16479
repeat_region /note="MIR repeat: matches 48. .262 of consensus"
16840. .17032
repeat_region /note="MER3 repeat: matches 2. .207 of consensus"
17472. .17501
misc_feature /note="15 copies 2 mer ac 100% conserved"
18839. .19765
/note="CpG island"
evidence=not_experimental
repeat_region /note="AluJo repeat: matches 1. .297 of consensus"
20385. .20649
repeat_region /note="FLAM.A repeat: matches 4. .116 of consensus"
20642. .20734
repeat_region /note="FLAM.A repeat: matches 4. .116 of consensus"
22760. .23085
repeat_region /note="AluSx repeat: matches 1. .306 of consensus"
23275. .23506
repeat_region /note="AluJb repeat: matches 1. .225 of consensus"
23540. .23584
repeat_region /note="L2 repeat: matches 2647. .2694 of consensus"
23590. .23671
repeat_region /note="AluJ/FLAM repeat: matches 1. .83 of consensus"
23714. .24013
repeat_region /note="AluSx repeat: matches 1. .297 of consensus"
24532. .24817
repeat_region /note="AluJb repeat: matches 17. .293 of consensus"
25033. .25166
repeat_region /note="FLAM.C repeat: matches 11. .143 of consensus"
25261. .25549
repeat_region /note="AluSc repeat: matches 1. .289 of consensus"
25675. .25877
repeat_region /note="MTLJ repeat: matches 2. .214 of consensus"
25998. .26083
repeat_region /note="Charliel repeat: matches 3. .92 of consensus"
27261. .27556
repeat_region /note="AluSx repeat: matches 3. .289 of consensus"
28123. .28175
repeat_region /note="L2 repeat: matches 2678. .2733 of consensus"
28395. .28708
repeat_region /note="AluSg repeat: matches 1. .310 of consensus"
29346. .29766
repeat_region /note="L2 repeat: matches 2085. .2448 of consensus"
29767. .30052
repeat_region /note="AluSg repeat: matches 1. .287 of consensus"
30053. .30345
repeat_region /note="L2 repeat: matches 2448. .2750 of consensus"
31191. .31486
repeat_region /note="AluSg repeat: matches 1. .296 of consensus"
31804. .31932
repeat_region /note="AluJb repeat: matches 1. .127 of consensus"
32054. .32075
repeat_region /note="11 copies 2 mer aa 100% conserved"
32503. .32780
repeat_region /note="AluJo repeat: matches 1. .288 of consensus"
32791. .32817
repeat_region /note="9 copies 3 mer tgt 96% conserved"
32817. .33099
repeat_region /note="AluSg repeat: matches 1. .282 of consensus"
33156. .33450
repeat_region /note="AluJb repeat: matches 6. .301 of consensus"
33747. .34055
repeat_region /note="AluYa5 repeat: matches 1. .309 of consensus"
34584. .34775
repeat_region /note="L1ME repeat: matches 5644. .5823 of consensus"
34786. .34930
```

```
repeat_region /note="FLAM.C repeat: matches 12. .142 of consensus"
34931. .35225
repeat_region /note="L1ME repeat: matches 5265. .5563 of consensus"
35257. .35727
repeat_region /note="L1ME repeat: matches 4762. .5233 of consensus"
35743. .36021
repeat_region /note="AluJb repeat: matches 1. .287 of consensus"
36340. .36488
repeat_region /note="L1PA15 repeat: matches 6009. .6157 of consensus"
37329. .37776
repeat_region /note="L1MC5 repeat: matches 7275. .7680 of consensus"
38067. .38121
repeat_region /note="L2 repeat: matches 2681. .2735 of consensus"
38346. .38564
repeat_region /note="L1MB5 repeat: matches 5921. .6142 of consensus"
39036. .39268
repeat_region /note="MIR repeat: matches 20. .262 of consensus"
39278. .39564
repeat_region /note="AluSc repeat: matches 1. .294 of consensus"
39575. .39657
repeat_region /note="AluJo/FRAM repeat: matches 206. .289 of consensus"
41653. .41957
repeat_region /note="AluJo repeat: matches 5. .293 of consensus"
42324. .42634
repeat_region /note="AluY repeat: matches 1. .311 of consensus"
43028. .43056

Query Match 34.7%; Score 816.8; DB 9; Length 143618;
Best Local Similarity 99.8%; Pred. No. le-177;
Matches 818; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1035 agtaccacaagaagcagcaaaaatgatgaaccacaagaagcacaacaaatggaatcaagaagaa 1094
|||||
Db 45496 AGTACCACAAAAGCAGCAAAAATGATGAACCAACAAAGCACAATAATGCAATCAAGAAG 45437
|||||
QY 1095 tcttctccttggaacttttaggactcttgaaataagtcgcaaaagagagagtagttttgcac 1154
|||||
Db 45436 TCTTCTCTCCTTTGACTTTTAGGACTTCTGAAATAAGTGCAGAAAGAGAGAGTAGTTTGCAC 45377
|||||
QY 1155 cctgcataaataagcactcttttgaacttcttgagctaaataacagttttgacaaaaat 1214
|||||
Db 45376 CTTGCTAAATCAAGCACTCTCTTTTGACTTCTGAGACTAAATACAGTTTGCACAAAAT 45317
|||||
QY 1215 gctgcacacacacatgaatggcagacaagaagcattcccaatagttgggagcctcttcag 1274
|||||
Db 45316 GCTGCACACACCATGAATGGCAGACAAAGGCATTTTCCAATAGTTGGGAGCCTCTTCAG 45257
|||||
QY 1275 aagcatcaaaagtttgattgggctctcttbtgttgaggagtgcttaattctaaacct 1334
|||||
Db 45256 AAGCATCAAAGTTTGATTGGGCTCTCTTTTGTGAGGGATGTTCTAATTCTAAACCT 45197
|||||
QY 1335 gtaaatgcaggaagaatatttttaattcaaaaggtgccataacacgagccaatacaact 1394
|||||
Db 45196 GTAAATGCAGCAGGAAGATATTTTAATCAAAAGGTGCCAATACACGGACCAATCAACT 45137
|||||
QY 1395 cctttgaaattgatacagcagagagaacaaagagaggtggcagcagaagaaaactttct 1454
|||||
Db 45136 CCTTTGAATTGATACAGCAGAGAGAAACCAAGGAGTGGCAGACGAAGGAAAACCTTTCT 45077
|||||
QY 1455 tatttggaaatctcaaccacatgattcttgttttttagatgcagggtcacaagaataag 1514
|||||
Db 45076 TATTTGGAATCTCAACCACATGATTTCTTTTGTAGAGATGCAGGCTCAAAAAGTAATG 45017
|||||
QY 1515 catgtttcttcagcagaactgaattattcactgceatattgactctaaacacccaatacact 1574
|||||
Db 45016 CATGTTTCTTCAGCAAACTGAATTAATCAGTCCCATATGACTCTAAACACCAAAATACGT 44957
|||||
QY 1575 aatgcctctaagttaaagcaccatgactctagtgtccttgggtgtatatattcttacacact 1634
|||||
Db 44956 AATGCCTCTAANTGAAGCACCATGACTCTAGTGTCTCTTGGTGTATATTCTTACATACCT 44897
|||||
QY 1635 tagtggaaaatcctctatttttctaatcatgctcccaagtggtaccagttctaaagtgtct 1694
|||||
```



```
Db 37 CCTGAGGAATATTCACAGAGGGTCCAGGCCATGAAGAGTCCTGACCACAAATGGGAGGA 96
Qy 107 ggagttgccaatgaattctgaagctgaaaggcaattctaccagatacaaggcagagcaa 166
Db 97 CAACTTCGCCGGGACTTCATCGCGTAAAGAGATGTGCTACCAAAATAGAACAGAAAA 156
Qy 167 aacctatacctaactgctggctggagaaatgcacaagaatcatcaagaaacacagataaagga 226
Db 157 GATATATCCACAGCCACTGAGAAAAAGAAATAATGTTAAAGAACACATACAGGA 216
Qy 227 tatattgcccattgattatagccgggtagaactatccctgataaacctctgtagagattc 286
Db 217 CATACTGCCAATTGATCACGCCGAGTTAAATTGACATTAAGACTCCTTCACAAGATTC 276
Qy 287 cagctacatcaatccaacttcattaaggagatttatggaccgaagcttatattggcac 346
Db 277 AGACTATATCAATCAAAATTTTAAAGGGCGCTATGGGCCAAAGCATATGTAGCAAC 336
Qy 347 ccagggtcctttatctacaacccctcctggactctctggaggatgattgggaatatagtg 406
Db 337 TCAAGGACCTTTAGCAATACAGTAAGATATTTTGGAGGATGCTATGGGAGTATATGT 396
Qy 407 ccttatactgttatggcatgcatggaatggaatgggaagaaagaaagtgtgagcgcta 466
Db 397 TGTGATCATTTGATGGCTGCCGAGAAATTTGAGATGGGAAGAAAAAATGTGAGCGCTA 456
Qy 467 ctgggctgagccagagagatgagctggaatttggcccttctctgtatcctgtgaagc 526
Db 457 TTGGCCTTTGATGGAGGAAGACCCCATTAACGTTTGCACCATTTAAAAATTTCTTGAGGA 516
Qy 527 tgaaaaaggaaatctgattataataacaggactctaaagttaagtccaatagtgaaac 586
Db 517 TGAACAAGCAAGACAGACTACTTCATCAGGACACTCTTACTTGAATTCAAATGAATC 576
Qy 587 tcgaactatctaccagtttctatacaagaattggccagaccatgatgtaccttcatctat 646
Db 577 TCGTAGGCTGTATCAGTTTCATATGTAAGTGAACCTGGCCAGACCATGATGTTCTTCATCAT 636
Qy 647 agacctattcttagctcatctgggatgtacgtgtttaccagagagatgacagtggtcc 706
Db 637 TGATCTTATCTGGACATGATGAAGCTTAATGAGGAATATCAAGAACATGAAGATGTCC 696
Qy 707 catatgcattcactgcagtcgctgtggaaggactgtgttattgtgtctattgttga 766
Db 697 TATTTGATTCATTGCAGTCAGCTGTGGAGAACAGAGTGCCATTTGTGCCA---TAGA 753
Qy 767 ttatacatgtagtctgctaaagatgggataattcctgagaaactcagtggtttcagttt 826
Db 754 TTATACGTGGAAATTTACTAAAAAGCTGGGAAAAATACCAGAGCAATTTAATGTATTAAATTT 813
Qy 827 gatccgggaaatgcggacacagagcccttcattagttcaaacgcaggagaaacaatatgaact 886
Db 814 AATACAGAAATAGAAACACAAAGGCATTCGCAAGTACAAACAAAGGACGAATATGAAC 873
Qy 887 ggtctacaatgctgtatttagaactatttaagagacagat 925
Db 874 TGTTCATAGACTATGTCACCACTGTTTGAAAAACAGCT 912

RESULT 12
G18337
LOCUS G18337 2938 bp DNA linear STS 28-SEP-1998
DEFINITION SWS1985 Eric D. Green Homo sapiens STS genomic, sequence tagged
site.
ACCESSION G18337
VERSION G18337.1 GI:1222794
KEYWORDS STS:
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2938)
AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE 2 (bases 1 to 2938)
AUTHORS Green,E.D.
TITLE Human chromosome 7 STS (1997)
JOURNAL Unpublished
COMMENT Synonyms: PTPN12
GDB: GDB:3754375
GDB_DSEG: PTPN12
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@hghri.nih.gov
Primer A: GTAAGAAATTTCTGCATGG
Primer B: TTTCTTAAAACTCCAGGG
STS size: 73
PCR Profile:
  presoak: 94 degrees C for 2.50 minute(s)
  Denaturation: 94 degrees C for 0.50 minute(s)
  Annealing: 45 degrees C for 0.75 minute(s)
  Polymerization: 72 degrees C for 0.75 minute(s)
  PCR Cycles: 35
  Thermal Cycler: PerkinElmer 9600

Protocol:
  Template: 30-100 ng
  Primer: each 2 uM
  dNTPs: each 100 uM
  Taq Polymerase: 0.05 units/ul
  Total Vol: 10 ul

Buffer:
  MgCl2: 1.5 mM
  KCl: 100 mM
  Tris-HCl: 10 mM
  NH4Cl: 5 mM
  pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: D13380 For additional information about the NHGRI chromosome 7 mapping project, see http://www.nhgri.nih.gov/DIR/STB/CHR7. Also see Genomics 11:548-64 (1991) [NUID=92128937].

FEATURES
    source
        1..2938
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /map="7"
        /clone_lib="Eric D. Green"
    gene
        1..2938
        /gene="PTPN12"
    STS
        2760..2832
        /gene="PTPN12"
        2760..2778
        /gene="PTPN12"
    primer_bind
        primer_bind complement(2815..2832)
        BASE COUNT 1004 a 556 c 577 g 801 t
        ORIGIN

Query Match 16.0% Score 376.6; DB 11; Length 2938;
Best Local Similarity 65.1%; Pred. No. 2.8e-76;
Matches 572; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

Qy 47 ccaagagaaattctgcagaagttcctggatgagcccaagcaagaaattactaaaga 106
Db 37 CCTGAGGAATATTCACAGAGGGTCCAGGCCATGAAGAGTCCTGACCACAAATGGGAGGA 96
```

QY	107	ggagtttgccaatgaatttctgaagctgaaaaggcaatctaccaagtacaaagcgagacaa	166
Db	97	CAACTTCGCGCGGACTTCATCGGTTAAGAAGATTGTCTACCAATATAGAACAAGAA	156
QY	167	aacctatctacaactgctgctgagaaatcccaagaatacaagaagaaacagataagga	226
Db	157	GATATATCCACAGCCACCTGGAGAAAAGAAAGAAATTTAAAAAGAACAGATACAAG	216
QY	227	tatttgccctatgattatagccgggtagaactatccctgataacctctgatgagattc	286
Db	217	CATACTGCCATTGTATCAGACGGGAGTTAAATTGACATTAAGACCTCTCAAGATT	276
QY	287	cagctacatcaatgccaacttcattaaaggagtttatgaccaccaagcgttatattgccac	346
Db	277	AGACTATATCAATGCAAAATTTATAAAGGCGCTCTATGGCCAAAGACATATGTAC	336
QY	347	caagggtcctttatctacaacctcctggactctctggaggtgatttgggaatatagtgt	406
Db	337	TCAAGGACCTTTAGCAAAATACAGTAATAGATTTTGGAGAGTGGTATGGGAGTAA	396
QY	407	ccttatcatgttatggcatgcatgagtatgaaatgggaagaaagtgagcgcta	466
Db	397	TGTGATCATGTGTAATGGCTCGGAGAAATTGAGATGGGAAGAAATATGTAGCGCTA	456
QY	467	ctgggctgagccaggagagatgcagctggaatttggccctctctctgtatcctctgaa	526
Db	457	TTGCCCTTTGTATGGAGAGACCCCAATAAGTTTGACCATTTAAATTTCTTGTGAG	516
QY	527	tgaaaaaagaaatctgtatataatcsggactctcaaaagttaaatcaatagtgaaac	586
Db	517	TGAACAAGCAAGAACAGACTACTTTCATCAGGACACTCTTACTTGAATTTCAAAAT	576
QY	587	tgaactatctaccagtttattacaagaattggccagaccatgatgtaaccttcatctat	646
Db	577	TCGTAGGCTGTATCAGTTTCAATTTATGTGAAGTGGCCAGACCATGATGTTCCIT	636
QY	647	agacctattctgagctcatctggatgatacgttattaccagagagatgacagtgttcc	706
Db	637	TGATTTCTTGTGACATGATAAGCTTAATGAGGAATATCAAGAACATGAAGATGTC	696
QY	707	catatgcattcaactgagctgctggtggaaggagctgtgtatttctgctattgttga	766
Db	697	TATTTCTATTTCAATGTCAGTGCAGGCTGTGGAAGAACAGTGCCATTTGTGCCA	753
QY	767	ttatcatgagtgctgctaaaagatggataattcctgagaactcagtggtttcagttt	826
Db	754	TTATACGTGGAATTTACTAAAGCTGGGAAATATACCAGAGGAATTTAATGTAATTT	813
QY	827	gatccgggaaatcggcacagagcgcttcattagtttcaaacgcaggaacaatatgaact	886
Db	814	AATACAAGAAATGAGAACAACAAGGCATTCGCAGTACAAACAAGAGCAATATCACT	873
QY	887	ggtctacaatgctgtattgaactatttaagagacagat	925
Db	874	TGTTCTAGAGCTATTGCCCAACTGTTTGAATAACAGCT	912
RESULT 13			
MMPTPPES			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL	Biochem. J.	308 (Pt 2),	425-432	(1995)
MEDLINE	95289971			
REFERENCE	2 (bases 1 to 2676)			
AUTHORS	Tremblay, M.L.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-MAY-1995)	M.L. Tremblay,	McGill University,	Dept of
COMMENT	Biochemistry,	3655 Drummond,	Montreal,	Quebec H3G 1Y6, CANADA
FEATURES	Related sequences:	X63440 and	S36169.	
source				
1..2676				
/organism="Mus musculus"				
/strain="BALB/c"				
/db_xref="taxon:10090"				
/chromosome="5"				
/map="region A3-B"				
/clone_lib="Lambda Zap II"				
/dev_stage="17.5 days"				
1..14				
/note="synthetic linker"				
330..2657				
/gene="mtp-pest"				
330..2657				
/gene="mtp-pest"				
/codon_start=1				
/product="protein-tyrosine-phosphatase"				
/protein_id="CAA60477.1"				
/db_xref="GI:804998"				
/db_xref="MGD:MG1:104673"				
/db_xref="SWISS-PROT:P35831"				
/translation="MEQVEILRRFIQRVOAMKSPDHNGEDNFARFMRRLRLSTKYRT				
EKIYPTATGKEENVKRKYDKILFDHSRVKLTLPQSDSDYINAFIKGVGPKA				
YVATOGPFRTVIDFWIMWYVNVIMVACREFEMGRKKERYEPLXGEDITPAF				
KISCNEQARTDYEITLLLEFQNSRRLYQPHYVNPDPHVPSSFDSTLDMISLMRK				
YQEHEDVPICHCSAGCGRTGAICADYTNLLKAGKIPERENVNLTQEMTQRHSA				
VOTRQYELVHRAIAQLFEKQLQIEHGAQIRGNEITITGTMVQSDSDRYHKPPL				
KPPTRSLVEGDAKEEILQPEPEHPVPIPTSPSAFTVTVMQSDSDRYHKPPL				
HMASPEQHPADLNRSYSDQWKSSESAIEHIDKLERNLSFEIKKVPLOQPKSFD				
GNTLLNRGHAIKIKSASSVSDRTSQBELSAGALKVDDVSONSCADCSAAHSHRAAE				
SSEESQNSHTPPRDCPLDKKKHVTWSLHGPENATVPVSDPKSDKSPDNHSOTLKT				
SSTPNSATAEAEADLTTEHNSSPLLKAPLSFTNPLHSDDDHSDGSGDGVTRNKT				
STASATVSPASAESACHRRVLPMSIARQEVAGTPHSCNEKDADYSESPPLPERTP				
ESFVLADMPVPEWHLEPNQSESEGLTTSNGNEKHADGAGIITEASADSPPAFSD				
KKDQITKSPAETVITDFGNRCQKPKGPRPPEMT"				
BASE COUNT	774 a	632 c	682 g	588 t
ORIGIN				
Query Match	15.7%;	Score	370.2;	DB 10; Length 2676;
Best Local Similarity	64.9%;	Pred. No.	8.3e-75;	
Matches	587;	Conservative	0;	Mismatches 303; Indels 15; Gaps 2;
QY	33	ctctgcagcatgacacaaagaaattctgcagaattcctgagatgagggcccaagaag	92	
Db	321	CGCGGAGGATGGAGCAAGTGGAGATCCCGAGGAGGTTTATCCAGAGGGTCCAGGCATG	380	
QY	93	aaattac-----taaaagaggagtttgcgaatgaatttctgaagctgaaaaag	140	
Db	381	AAAGATCCGGATCACAAATGGGAGGACAACTTCGCCCGGGACTTCATGCGATTGAGA	440	
QY	141	caatctaccaagtacaaaggcagacaaacctatctcactgaactggtgagatgccaag	200	
Db	441	TTGCTTACCAATATAGAACAGAAAAGATTATCCCAAGCCACTGAGAAAAAGAA	500	
QY	201	aatatcaagaaaaacagatataaggatattttgcccctatgattatagccgggtagaact	260	
Db	501	AATGTTAAAAAAGAACAGATATAAGGACATACTGCCATTTGATCACAGCCCGATTAG	560	
QY	261	tcctgtataacctctgatgaggtatccagctacatcaatgccaaacttcattagggagtt	320	
Db	561	ACTTGAAGACTCCATCCCAAGATTCAGATTATATCAATGCAATTTTATTAGGGTGTG	620	
QY	321	tatggacccaaggcttatattgcccacccagggttcctttatctacaaacctctcgtgacttc	380	

Db 846 ACAGTGCTATTGTGGCCA--TCGACTACACGTGGAACCTTACTGAAAGCAGGGAATAAT 902
Qy 801 cctggaactcagtggtttcaggttgatccggaataatcggaacacagagcccttcatta 860
Db 903 CCAGAGGAATTAATGTATTAATTAACAAGAAATGAGAACAACAAGGCACCTCTGCA 962
Qy 861 gttaaacgcaggaacaatatgaactggctctacaaatgctgtattagaactatttaagaga 920
Db 963 GTACAAACGAAGGACGATGATGAACCTCTCCATAGAGCTATTGCCCACTGTTGAAAAA 1022
Qy 921 cagat 925
Db 1023 CAGCT 1027

RESULT 15
AX306312 AX306312 2983 bp DNA linear PAT 11-DEC-2001
LOCUS Sequence 1063 from Patent WO0188188.
DEFINITION AX306312
ACCESSION AX306312
VERSION AX306312.1 GI:17645552
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1063 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
source 1..2983
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 923 a 634 c 681 g 745 t
ORIGIN

Query Match 15.6%; Score 368; DB 6; Length 2983;
Best Local Similarity 64.9%; Pred. No. 2.6e-74;
Matches 583; Conservative 0; Mismatches 300; Indels 15; Gaps 2;

Qy 33 cctctcagcatggaccagaagaaattctgcagaagattcctggatgagggcccaagcaag 92
Db 43 CGCGGGAGGATGGAGCAAGTGAGATCCTGAGGAGGTTTCATCCAGAGGGTCCAGGCCATG 102
Qy 93 aaaatac-----taaagagaggtttgccaatgaattctcgaagctgaaaagg 140
Db 103 AAGAGTCCGGATCACAAATGGGGAGGACAACTTCGCCGGGACTTCATCGATTGAGAAGA 162
Qy 141 caatcaccaagtacaaggcagacaaaacctatctcacaactgtgctgagaatcccaag 200
Db 163 TTGCTTACCAATATAGAACAGAAAAGATTTATCCACAGCCACTGGGAGAAAAAGAA 222
Qy 201 aatatcaagaaaaacagataaagatattttgcctctgattatagcgggtagaacta 260
Db 223 AATGTTAAAAGACAGATATAAGGACATACCTGCCATTGTATCAGCCGCGGTTAAGTTG 282
Qy 261 tccctgataaacctctgatgaggtatccacgtcacatcaatgccaaacttcattaaaggagtt 320
Db 283 ACTTTGAAGACTCCATCCCAAGATTTCAGATTATATCAATGCAAAATTTATTAAAGGTGTG 342
Qy 321 tatggaccacaagcttatattgccaccagggtcctttatctacaacccctcctggacttc 380
Db 343 TATGGGCCAAAAGCATATGTGGCAACCCCAAGGGCCCTTTCCGGAATACAGTCTATGACTTC 402
Qy 381 tggagagatttgggaataatagttcccttcattctgttattggtcatgcatgaggtatgaa 440
Db 403 TGGAGGATGATATGGAGGTATTAATGTTGTGATGATCGTGCATGCCCTGTCGAGAAATTTGAG 462
Qy 441 atgggaaagaaaaagtgtgagcgtactggtggctgagccaggagagatgcagctggaattt 500

Db 463 ATGGGAAGAAAAAGTGTGAGCGCTACTTGGCTTTTGTATGGAGAAGATCCTTATAACATTT 522
Qy 501 ggccttctctgtatcctctgaaagctgaaaaaaggaataatctgattatataatcaggagact 560
Db 523 GCACCATTTAAAATTTCTGTGAAANTGAACAAGCAAGAACCGACTACTTTCATCCGAACA 582
Qy 561 ctaaaagttaagtccaatagtgaaactcgaactatctaccaggttttcattacaagaattgg 620
Db 583 CTTTACTTGAATTTCAAAATGAATCCCGTCGGCTCTATCATCAGTTTCATTACGTGAACCTGG 642
Qy 621 ccagaccatgatgtaccttcattctatagaccctattcttgagctcatctctggatgagctg 680
Db 643 CCAGACCATGATGTTCCCTCGCTTCATTTGATTTCTATTCTGGACATGATTAAGCTTAATGAGG 702
Qy 681 tgttaccagaaggatgacagtggttcccatatgcattcactcagtgctgctgctggaagg 740
Db 703 AAATACCAGAACAATGAAGATGTGCCTATTGTATTTCAATTCAGTGTCTGGCTGTGGACGA 762
Qy 741 actggtgttattgtgctattgttgattatcacatggtattgtctaaaaagatgggataatt 800
Db 763 ACAGGTGCTATTGTGCCA---TAGATTACACGTGGAACCTTACTGAAAGCAGGGAAAAAT 819
Qy 801 cctgagaacttcagtggttttccagtttgatccgggaaatcgcggaacacagagagcccttcatta 860
Db 820 CCAGAGGAATTTAATGTATTTAATTAACAAGAAATGAGAACAACAGAGGCACCTCGGCA 879
Qy 861 gtccaacgcaggaacaataatgaactggtctacaaatgctgtattagaactatttaaga 918
Db 880 GTACAAACAAAGGAGCAGTATGAACCTTGTTCATAGGGCTATTGCTCAACTGTTTGAAA 937

Search completed: May 26, 2002, 10:35:23
Job time: 12938 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 10:42:10 ; Search time 56.14 Seconds
(without alignments)
1184.429 Million cell updates/sec.

Title: US-09-600-358A-4
Perfect score: 3615
Sequence: 1 MDQREILQKFLDEAQQSKIT.....RPSKSVKLRSPKSGKNFSQL 692

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2356.5	65.2	802	1 B44390	protein-tyrosine-p
2	1157	32.0	780	1 JCL138	protein-tyrosine-p
3	1127.5	31.2	775	2 S53345	protein-tyrosine-p
4	1104.5	30.6	773	1 JH0609	protein-tyrosine-p
5	1042	28.8	382	1 S48748	protein-tyrosine-p
6	514.5	14.2	1304	1 A46546	leukocyte common a
7	513	14.2	1304	2 S68700	HPRP beta-like tyr
8	508	14.1	1337	1 I38670	protein-tyrosine-p
9	507	14.0	1291	1 A28334	protein-tyrosine-p
10	493	13.6	1273	1 TDRTLT	protein-tyrosine-p
11	490.5	13.6	1440	2 JC6312	leukocyte common a
12	485	13.4	595	1 S20825	protein-tyrosine-p
13	485	13.4	1442	1 B48148	protein-tyrosine-p
14	483.5	13.4	593	2 JC5167	protein-tyrosine-p
15	483.5	13.4	1200	2 T43148	protein-tyrosine-p
16	479	13.3	1452	1 S17670	probable protein-t
17	478.5	13.2	1457	1 A48066	protein-tyrosine-p
18	478	13.2	2314	1 A46151	protein-tyrosine-p
19	477	13.2	1452	1 S17669	protein-tyrosine-p
20	476	13.2	1445	1 A48148	protein-tyrosine-p
21	475.5	13.2	1237	2 A54080	protein-tyrosine-p
22	473	13.1	593	1 JN0805	protein-tyrosine-p
23	471.5	13.0	597	1 A53593	protein-tyrosine-p
24	471	13.0	1912	2 A56178	protein-tyrosine-p
25	470	13.0	595	1 A55651	protein-tyrosine-p
26	469.5	13.0	1691	1 D54689	protein-tyrosine-p
27	469.5	13.0	1894	2 C54689	protein-tyrosine-p
28	469	13.0	595	1 A44390	protein-tyrosine-p
29	467.5	12.9	582	2 A57068	protein-tyrosine-p

ALIGNMENTS

RESULT 1

B44390

protein-tyrosine-phosphatase (EC 3.1.3.48) PEP, nonreceptor type 8 - mouse
N:Alternate names: protein-tyrosine-phosphatase PEP
C:Species: Mus musculus (house mouse)
C>Date: 03-May-1994 #sequence_revision 26-May-1994 #text_change 11-Jun-1999
C:Accession: B44390; S71952; S27876
R:Matthews, R.J.; Bowne, D.B.; Flores, E.; Thomas, M.L.
Mol. Cell. Biol. 12, 2396-2405, 1992
A:Title: Characterization of hematopoietic intracellular protein tyrosine phosphatase
A: and threonine-rich sequences.
A:Reference number: A44390; MUID:92236615
A:Accession: B44390

A:Molecule type: mRNA

A:Residues: 1-802 <MAT>

A:Cross-references: GB:M90388; NID:g200522; PIDN:AAA39994.1; PID:g200523

R:Cloutier, J.F.; Veillette, A.

EMBO J. 15, 4909-4918, 1996

A:Title: Association of inhibitory tyrosine protein kinase p50(csk) with protein tyro

A:Reference number: S71952; MUID:97045099

A:Accession: S71952

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 495-789 <CLO>

C:Comment: This protein is found primarily in hematopoietic tissues.

C:Genetics:

A:Gene: 70zpep

C:Complex: physically associates with inhibitory tyrosine protein kinase Csk; interac

C:Function:

A:Description: probably an effector and/or regulator of tyrosine protein kinase csk 1
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 8; protein-tyrosine-pho
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phospho
F:54-278/Domain: protein-tyrosine-phosphatase homology <PTP>
F:497-802/Region: glutamic acid/proline/serine/threonine-rich
F:613-621/Region: proline-rich
F:688-695/Region: proline-rich

F:227/Active site: Cys (phosphocysteine intermediate) #status predicted

F:233/Binding site: substrate phosphate (Arg) #status predicted

Query Match 65.2%; Score 2356.5; DB 1; Length 802;

Best Local Similarity 68.5%; Pred. No. 5.2e-139;

Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

Qy 1 MDQREILQKFLDEAQQSKITKEEFANFLKRSQPKYKADKTYPTTVAENAKNKKNY 60

Db 1 MDQREILQKFLDEAQQSKITKEEFANFLKRSQPKYKADKTYPTTVAENAKNKKNY 60

Qy 61 KDILPYDYSRVLSLTSDSSYINANFKGVYGPAYATQGPLSTTLLDFWRMIWEY 120

Db 61 KDILPYDHSVLVLSLTSDSSYINASFKGVYGPAYATQGPLSTTLLDFWRMIWEY 120

QY 632 AGEFSNPVKSLSSAVKVICTGLEWG-----GTSBPKKFDDSVILRPSKSVKLRSKPSG 686
Db 691 -----HNTP-----VRSSELSQSRSEKQSEG---LITSEKCDHPAGG 730
QY 687 KNF 689
Db 731 IHY 733
RESULT 3
S55345
protein-tyrosine-phosphatase (EC 3.1.3.48) PTP-PEST - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Oct-1995 #sequence_revision 01-Nov-1996 #text_change 22-Jun-1999
C:Accession: S55345; S54261: I48666
R:Charest, A.; Wagner, J.; Shen, S.H.; Tremblay, M.L.
Biochem. J. 308, 423-432, 1995
A:Title: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein tyrosine P
A:Reference number: I48666; MUID:95289971
A:Accession: S55345
A:Molecule type: DNA
A:Residues: 1-775 <CHA>
A:Cross-references: EMBL:X86781; NID:g804997; PIDN:CAA60477.1; PID:g804998
R:Charest, A.; Wagner, J.; Shen, S.H.; Tremblay, M.L.
submitted to the EMBL Data Library, May 1995
A:Description: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein tyro
A:Reference number: S54261
A:Accession: S54261
A:Molecule type: DNA
A:Residues: 1-127,'M',129-309,'R',311-775 <CHW>
A:Cross-references: EMBL:X86781; NID:g804997; PIDN:CAA60477.1; PID:g804998
A:Experimental source: strain BALB/c
C:Genetics:
A:Map position: 5A3-B
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatas
F:58-282/Domain: protein-tyrosine-phosphatase homology <PTP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted
Query Match 31.2%; Score 1127.5; DB 2; Length 775;
Best Local Similarity 36.6%; Pred. No. 1.9e-62;
Matches 280; Conservative 109; Mismatches 166; Indels 209; Gaps 28;
QY 1 MQOREILOKFLDEAQSCKIT-----KSEFANEFLKLRQSTKYKADKYPTTYVAENAKNIK 56
Db 1 MEQVEILRRFIQRVQAMKSPDHNGEDNFARDFMLRLRLSTKYTEKIYPTATGEKEENVK 60
QY 57 KNYKDIPLDYSRVLSLITSDSSYINANFIKGVYGPXATGATQGLSTLLDFWRM 116
Db 61 KNYKDIPLFDHSRVKLTLPQSDSDYINANFIKGVYGPXAVATQGFPRNTVIDFWRM 120
QY 117 IWEYSVLIIIVMACYEMGKKKERYWAEFGEMQLEFGFPFSVCEAEKRSYDIIRTLKV 176
Db 121 IWEYNVVIIVMACREFEMGRKCYRWPLYGEDPITFAPEFKISCENEQAQTDYFIRTL 180
QY 177 KFNSETRTIYQPHYKNPDHVPSSIDPILIELIWDVRCQEDDSVPICHCSAGCGRTGV 236
Db 181 EFQNESRRLYQFHYVNMVDHVPSSFDSDILDMISLMRKYOEHEDVDVPICHCSAGCGRTGA 240
QY 237 ICAIVDYTWMLLGIIPNFVSFLIREMRTORPSLVOTEOQYELVYNVAVLELFFKROMD 296
Db 241 ICAI-DYTNLLKAGKIPEEFNVFNLIQEMRTORHSAVQTKQYELVHRIAQLFQKQ 299
QY 297 V-----IRDKHSGTESQAKHCEPKNHTLOADYSNLPKSTTKAAKMMNQRTKME 348
Db 300 LYIEHGAQKIADGNEITTGTGMWSSIDSE-----KQDSPPPKPPPTRS-----CLVEGDAKEE 351
QY 349 IKE-----SSSFDFRIS-----EISAKEELVLPKAKSTSTF 379
Db 352 ILQPPPEHPVPPIILTPSPSFAFTVTVMQDSDRYHPKPYLHMASPEQ---HPA----- 402

QY 380 DFLELNYSFQKADNTTKMW-----QTKAPFIVGEPLQK-HOSLDLGLSLLF 423
Db 403 ----DLARSYDKSAD---QWCKSESAIEHIDKKLERNLSEIKKVPLOEGPKSFDGNTLLN 456
QY 424 EGCSNKPVNAAGRYFNKSVPIITRTKSTPFEL-----IQORETKEV----- 454
Db 457 RG--HAIKIKSA-----SSSVVDRT--SKPOELSAGALKVDDVDSQNSCADCSAAHSHRAAE 508
QY 465 DSKENFYLESDPHSCFVEMQAQKVMHYSSAELNYSL-----PYDSK 507
Db 509 SSEESQSNHTPPRPDC---LPLDKKGHYT-----WSLHGPNATVPVPSDCKGSPDNHS 560
QY 508 HQIRNASN-----VKHDSALGVYSYIPLVENP-----YFSSWPPSGTS 547
Db 561 QTLKTVSSTPNSTAEAEADLTEHNSS-----PLLKAPLSFTNPLHSDDDHSDGGS 612
QY 548 SKMSLDLPKQDQGVFPFSSLLPSTSLFSYNSHSLNSLSPNSSLNQESAV----- 603
Db 613 S-----DGAIV-----TRNKTSISTASAT-VSPASSAESACHRRV 645
QY 604 -----LATAPR-----IDDEIPPLPVRTPESEFIVVE 630
Db 646 LPMISARQEVAGTPHSGAEKADADVSESPPLPERTPESEFVLAD 689
RESULT 4
JH0609
protein-tyrosine-phosphatase (EC 3.1.3.48) P19 - mouse
N:Alternate names: protein-tyrosine-phosphatase PPTV43
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0609; PS0365; PS0369; PS0366; G61180
R:den Hertog, J.; Pals, C.E.G.M.; Jonk, L.J.C.; Kruijer, W.
Biochem. Biophys. Res. Commun. 184, 1241-1249, 1992
A:Title: Differential expression of a novel murine non-receptor protein tyrosine phos
A:Reference number: JH0609; MUID:92272714
A:Accession: JH0609
A:Molecule type: mRNA
A:Residues: 1-773 <DEN>
A:Cross-references: GB:X63440; GB:S36169; NID:g416181; PIDN:CAA45037.1; PID:g416182
A:Experimental source: embryonic carcinoma cell, P19 cell
A:Accession: PS0365
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-110,'G',112-118,'S',120,'T',122 <DE2>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP33
A:Accession: PS0369
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'G',93-109,'LG',112-120,'T',122 <DE3>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP59
A:Accession: PS0366
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 88-91,'KY',94-109,'LA',112-118,'S',120-122 <DE4>
A:Experimental source: embryonic carcinoma cell, P19 cell, clone PTP42
R:Yi, T.; Cleveland, J.L.; Ihle, J.N.
Blood 78, 2222-2228, 1991
A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells
A:Reference number: A61180; MUID:92032882
A:Accession: G61180
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 124-127,'I',129-229 <YIA>
C:Comment: This protein is located in the cytoplasm.
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-ph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosph
F:55-299/Domain: phosphatase catalytic domain #status predicted <PCD>
F:58-282/Domain: protein-tyrosine-phosphatase homology <PTP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted

Query Match 30.6%; Score 1104.5; DB 1; Length 773;
Best Local Similarity 35.6%; Pred. No. 5e-61;
Matches 273; Conservative 103; Mismatches 175; Indels 215; Gaps 26;

Qy 1 MDQREILQKFLDEAQSCKIT----KEEFANFELKLRKSTYKADKTYPTTVAENAKNIK 56
Db 1 MEQVEILRRFQORVOAMKSPDHNGEDNFARDFMRRLRSTYKTEKIYPTATGEKEENVK 60
Qy 57 KNRKIDILPDYSRVELSLITSDSSYINANFKGVGPKAYIATOGPLSTLLDFWRM 116
Db 61 KNRKIDILPFDHSRVKLLTKTPSQSDSYINANFKGVGPKAYATQFFRTVIDFWRM 120
Qy 117 IWEYSLIIVMACMEYEMGKKKERYWAPGEMQLEFQFVSCEAEKRSYIIRTLLK 176
Db 121 IWEYNVIVMACREFEMGRKCRYWPLYGDPITFAFKTSCNEQARTDYFIIRTL 180
Qy 177 KNSETRITQFYHYKNWPDHVPSSIDPILFLWVRCYQEDDSVPICHSAGCGRTGV 236
Db 181 EFQNSRLRYQHYVNWPDHVPSSIDPILFLWVRCYQEDDSVPICHSAGCGRTGA 240
Qy 237 ICAIVDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOQYELVYNVLELFRKQMD 296
Db 241 ICAI-DYTWMLLKAGKIPEEFNVFNLIQEMRTQPSLVQTOQYELVYNVLELFRKQMD 299
Qy 297 VI-----RDKHS-----GTSQAKHCPIE 315
Db 300 CMKMEHRRSVVMKLPLELNSVPLIARRDLAKAAADSKLPCRRGCGQGRNTATRTSPG 359
Qy 316 KNTHTQADSYSNLPKSTTKAAM-----MNQORTKWEIKESSPDRFISE 361
Db 360 AHPDAITSFS--LPNVTVTCVTGTGTSQOCCTWPHOSNTQPTSTE-----AMINORT-- 411
Qy 362 ISAKELVLHPAKSTSFDFELNYSFKDNADTTMKWOTKAPPIVGEPLQKHQDLGLSL 421
Db 412 -----NGAKSESALEHI-----DKLERNLSPELKKVPLQEGP-----KSFQGNL 452
Qy 422 LFEGCSKNPVNAAGRYNSKVPITRTKSTPEL-----IQORETKEY----- 464
Db 453 LNRG--HAIKIKSA-----SSSVVDRT--SKQELSGALGVDDVDSQNSCADCSAAHSHRA 504
Qy 465 --DSKENFSLYESQPHDSFCFVEMQAQKVMYHVSASLNYSL-----PYD 505
Db 505 ARSESOSNSHTPPRPDC--LPDLKKGHVY----WSLHGPENATVPVPSDQKSPDN 556
Qy 506 SRHQIRNASN-----VKHHDSSALGVSYIPLVNP-----YFSSWPPSG 545
Db 557 HSQTLKTVSSTNSTAEAEADLTHEHNS-----PLKAPLSFTNPLHSDMDHSDG 608
Qy 546 TSKMSLDLPEKQDGTVPFSSLLPTSSLSFYSYNSHSLNSPTNSSLNQSSAV-- 603
Db 609 GSS-----DGAV-----TRNKTSTISAT--VSPASSAESACHR 641
Qy 604 -----LATAPR-----IDDEIPLPPLVPRTPEFIVVE 630
Db 642 RVLPMIARQEVAGTHSGAEKADADVSESPPLPPTPEFVLAD 687

RESULT 5
S48748
protein-tyrosine-phosphatase (EC 3.1.3.48), probable nonreceptor type 12 splice form -
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S48748
R:Moriyama, T.; Kawanishi, S.; Inoue, T.; Imai, E.; Kaneko, T.; Xia, C.; Takenaka, M.; N
FEBS Lett. 353: 305-308, 1994
A:Title: cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPTP) from rat kidne
A:Reference number: S48748; MUID:95046282
A:Accession: S48748
A:Molecule type: mRNA
A:Residues: 1-382 <MOR>
A:Cross-references: GB:D38072; NID:9567262; PIDN:BAA07266.1; PID:G699627
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 12; protein-tyrosine-phosph
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase

F:58-282/Domain: protein-tyrosine-phosphatase homology <PPP>
F:231/Active site: Cys (phosphocysteine intermediate) #status predicted
F:237/Binding site: substrate phosphate (Arg) #status predicted

Query Match 28.8%; Score 1042; DB 1; Length 382;
Best Local Similarity 53.5%; Pred. No. 1.4e-57;
Matches 201; Conservative 62; Mismatches 89; Indels 24; Gaps 5;
Qy 1 MDQREILQKFLDEAQSCKIT----KEEFANFELKLRKSTYKADKTYPTTVAENAKNIK 56
Db 1 MEQVEILRRFQORVOAMKSPDHNGEDNFARDFMRRLRSTYKTEKIYPTATGEKEENVK 60
Qy 57 KNRKIDILPDYSRVELSLITSDSSYINANFKGVGPKAYIATOGPLSTLLDFWRM 116
Db 61 KNRKIDILPFDHSRVKLLTKTPSQSDSYINANFKGVGPRAYVATOGPLANTVIDFWRM 120
Qy 117 IWEYSLIIVMACMEYEMGKKKERYWAPGEMQLEFQFVSCEAEKRSYIIRTLLK 176
Db 121 IWEYNVIVMACREFEMGRKCRYWPLYGDPITFAFKTSCNEQARTDYFIIRTL 180
Qy 177 KNSETRITQFYHYKNWPDHVPSSIDPILFLWVRCYQEDDSVPICHSAGCGRTGV 236
Db 181 EFQNSRLRYQHYVNWPDHVPSSIDPILFLWVRCYQEDDSVPICHSAGCGRTGA 240
Qy 237 ICAIVDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOQYELVYNVLELFRKQMD 296
Db 241 ICAI-DYTWMLLKAGKIPEEFNVFNLIQEMRTQPSLVQTOQYELVYNVLELFRKQMD 299
Qy 297 VIRKHSSTESQAKHCIPKNTHTQADSYSNLPKSTTKAAMNQORTKWEIKESSSFD 356
Db 300 -LYETHGAKITDGNELSTGNMVSIDSEKQDSDPP-----KPPRTSCLVEG----- 346
Qy 357 FRTSEISAKEELVHP 372
Db 347 -----DAKEELQPP 356

RESULT 6
A46546
leukocyte common antigen long splice form precursor - human
N:Alternate names: C045; protein-tyrosine-phosphatase, receptor type c; T200 glycopro
N:Contents: leukocyte common antigen intermediate splice form; leukocyte common anti
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A46546; B46546; C46546; A29449; B29449; I57658
R:Streuli, M.; Hall, L.R.; Saga, Y.; Schlossman, S.F.; Saito, H.
J. Exp. Med. 166, 1548-1566, 1987
A:Title: Differential usage of three exons generates at least five different mRNAs en
A:Reference number: A46546; MUID:88061067
A:Accession: A46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1304 <STR>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.6/2
A:Accession: B46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-32,99-264 <ST2>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.111 and clone LCA.260
A:Accession: C46546
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-31,193-264 <ST3>
A:Cross-references: GB:Y00638
A:Experimental source: clone LCA.1
R:Ralph, S.J.; Thomas, M.L.; Morton, C.C.; Trowbridge, I.S.
EMBO J. 6, 1251-1257, 1987
A:Title: Structural variants of human T200 glycoprotein (leukocyte-common antigen).
A:Reference number: A91066; MUID:87275816
A:Accession: A29449

A:Molecule type: mRNA
A:Residues: 1-31,193-649,'L',651-869,'G',871-872,'A',874-1206,'P',1208-1304 <R>
A:Cross-references: GB:Y00062; NID:934275; PIDN:CAA68269.1; PID:934276
A:Experimental source: clones PHLC-1 and lambdaHLG1
A:Accession: B29449
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 32-192 <RA>
A:Experimental source: clone HLC-2
R:Tsai, A.Y.; Streuli, M.; Saito, H.
Mol. Cell. Biol. 9, 4530-4555, 1989
A:Title: Integrity of the exon 6 sequence is essential for tissue-specific alternative splicing
A:Reference number: I57658; MUID:90066468
A:Accession: I57658
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 146-192 <RES>
A:Cross-references: GB:M29253; NID:g187020; PIDN:AAA59497.1; PID:9553521
C:Genetics:
A:Gene: GDB:PTPRC; CD45
A:Cross-references: GDB:119768; OMIM:151460
A:Map position: lq31-lq32
C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain homolog
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd
F:594-1235/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:575-899/Domain: protein-tyrosine-phosphatase homology <PTP>
F:851/Active site: Cys (phosphocysteine intermediate) #status predicted
F:857/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.2%; Score 514.5; DB 1; Length 1304;
Best Local Similarity 28.2%; Pred. No. 6.3e-24;
Matches 166; Conservative 86; Mismatches 241; Indels 95; Gaps 19;

QY 2 DQREILQK-----FLDAQSKKITKE--EPANFELKLKROSTIKYKADTKY 44
DB 612 EQGELVERDEKQLMNVPEITHADILLETYYKRIKTIADGRPLAEFQSPRVFSK-----F 665

QY 45 PTTVAENAKIKNRYKIDLPYDSRVLSLTSDSSVINANFIKVGYPKAYIATQGLSTLLDFWMIWSEYSLII 125
DB 666 PIKARKPFQONKRVVDILPYDNRYVSEIDNGDAGSNYINASYIDGFEKPKYIAAOG 725

QY 105 PLSTLLDFWMTWEXSVLIIVACMEYEMGKKKERYNAEPGEMOLEFGFVSVCSEAEK 164
DB 726 PRETVDDFWMTWEQKATVIVWTRCEGNRNKCAEYWFPSMEGTRAFGDVVVKINQHK 795

QY 165 RKSDYIIRTLKV---KFNSETRTIYQHYKNWPDHVPSSIDPILFELIWDVRCYQEDDSV 221
DB 786 RCPDYIIQKLNIVNKKKEKATGREVTHIQFTSWPDHGVDPDPLHLLKLRNRVNAFNSFFSG 845

QY 222 PICIHCSAGCGRTGVCIAVDYTWMLLKDGIIPENFSVFLIREMRTQPSLVQTOGVELVNAVLELFKROMD 296
DB 846 PIYVHCSAGVGRGTGIIGID-----AMLEDETENKVDVYGVYVVKLRQRORCLMVQVEAQYI 901

QY 282 LVYNVAVLEL--FKRQMDVIRKHSGETSOAKHICPEKNHTLOADSYSPNLPK-STYKAAK 338
DB 902 LIHQALVEYVQFETEVNLSELPYLNHMKKRDPPSPLEAB--FQRLPSTRSWRTQH 959

QY 339 MNAQRTKMEIKESSDFRTSEISAKEELV-----HPAKSSTFDLELNYSFQKNAD 393
DB 960 IGNOEENKSNRNSNVPIDYDNRVPLKHELEMSKESHDSDSDSDSDSEEPSESKYNAS 1019

QY 394 TTKW-WOTKAPPIVGEPLQKHSIDJGSLF-----BGCNSKRPVNA 434
DB 1020 FIMSYWKPENYIAAQQPL-KETIGDEFWQMIFQRKVKIVMTELKHLQDQEIACAQ---YWG 1075

QY 435 AGR--YFNKSVPIRTR-KSTPEEL--IQQRETKEVDSKENFSY-----LESQPHD- 479
DB 1076 EGKQTYGDIEVDLTDKDSSTYTLRVPFLRHSKRKDSRTVYQYTNWSVEQLPAEPKEL 1135

QY 480 SCFVEMQAQKVMHVSAAELNYSFLPYDSKKHQIRNASNVKHHDSALGVY 527
DB 1136 ISMIQVVKQKLPQKNSSEGN-----KHKSTPLLH 1166

RESULT 7
S68700
HPTP beta-like tyrosine phosphatase precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68700
R:Kuramochi, S.; Matsuda, S.; Matsuda, Y.; Saitoh, T.; Ohsugi, M.; Yamamoto, T.
FEBS Lett. 378, 7-14, 1996
A:Title: Molecular cloning and characterization of Byp, a murine receptor-type tyrosi
A:Reference number: S68700; MUID:96140699
A:Accession: S68700
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1238 <KUR>
A:Cross-references: GB:D45212; NID:gl208432; PIDN:BAA08146.1; PID:gl208433
C:Genetics:
A:Map position: 2E1-2
A:Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III re
C:Keywords: signal sequence #status predicted <SIG>
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1238/Product: HPTP beta-like tyrosine phosphatase #status predicted <MAT>
F:267-347/Domain: fibronectin type III repeat homology <3FR>
F:966-1188/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1140/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1146/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.2%; Score 513; DB 2; Length 1238;
Best Local Similarity 39.0%; Pred. No. 7.2e-24;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

QY 15 QSKKIKKEEFANFELKLKROST-----KYKADK-----TYPTTVAENAKIKNRYKIDLP 65
DB 918 KSLIRLVENFEAYFKKQADSNCGFAEYEDLKLIGLSLKPYKTAETAEAKNRNNVLP 977

QY 66 YDSRVLSLTSDSSVINANFIKVGYPKAYIATQGLSTLLDFWMIWSEYSLII 125
DB 978 YDISRVLSVQTHSTD--DYINANMPGYHSHKDFIATQGLPNTLKDFFWMIWSEYSLII 1036

QY 126 VMACMEYEMGKKKERYNAEPGEMOLEFGFVSVCSEAEKRSYIIRTLKV--FNSETR 183
DB 1037 VMLTKCQVEQRTKCEEW--PSKQADYGDITVAMTSEVVLPEWTIRDFVVKMNQNSH 1094

QY 184 TIYQHYKNWPDHVPSSIDPILFELIWDVRCY--QEDDSVPICHCSAGCGRTGVCIAIV 241
DB 1095 PLRQFHFTSWPDHGVDPDPLHLLKLRNRVNAFNSFFSG 1154

QY 242 DYTWMLLKDGIIIPENFSVFLIREMRTQPSLVQTOGVELVNAVLELFKROMD 296
DB 1155 RLIIQIENENTV-----DVYGIYDLRMHRPLMVQTDQYVFLNQCVLDIIRAQND 1205

RESULT 8
I38670
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type J precursor - human
N:Alternate names: density enhanced phosphatase-1 (DEP-1); protein-tyrosine phosphata
C:Species: Homo sapiens (man)
C:Date: 01-Mar-1996 #sequence_revision 08-Mar-1996 #text_change 22-Jun-1999
C:Accession: I38670; I52599
R:Ostman, A.; Yang, Q.; Tonks, N.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 9680-9684, 1994
A:Title: Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhance
A:Reference number: I38670; MUID:95024024
A:Accession: I38670
A:Molecule type: mRNA
A:Residues: 1-1337 <RES>
A:Cross-references: EMBL:U10886; NID:g558754; PID:g558755
A:Experimental source: HeLa cells
R:Honda, H.; Inazawa, J.; Nishida, J.; Yazaki, Y.; Hirai, H.
Blood 84, 4186-4194, 1994
A:Title: Molecular cloning, characterization, and chromosomal localization of a novel

A:Reference number: I52599; MUID:95086212
A:Accession: I52599
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-216, 'LTCVRKAA', 225-260, 'G', 262-285, 'GTEGLDASNTSERSA', 302, 'S', 304, 'TAPVHDE'
A:Cross-references: GB:D37781; NID:9633072; PIDN:BA07035.1; PID:9633073
C:Comment: Enhanced expression of this protein with increasing cell density suggests a role in cell growth.
C:Genetics:
A:Gene: GDB:PTPRJ
A:Cross-references: GDB:385040; OMIM:600925
A:Map position: 19q13.4-19q13.4
C:Function:
A:Description: catalyzes hydrolysis of peptidyl-phosphotyrosine to peptidyl-tyrosine and phosphate.
C:Superfamily: protein-tyrosine-phosphatase, receptor type J; fibronectin type III repeat
C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-133/Product: protein-tyrosine-phosphatase, receptor type J #status predicted <MAT>
F:118-197/Domain: fibronectin type III repeat homology <3FNC>
F:206-283/Domain: fibronectin type III repeat homology <3FNB>
F:284-356/Domain: fibronectin type III repeat homology #status atypical <3FNC>
F:365-445/Domain: fibronectin type III repeat homology <3FND>
F:453-530/Domain: fibronectin type III repeat homology <3FNE>
F:539-617/Domain: fibronectin type III repeat homology <3FNF>
F:720-804/Domain: fibronectin type III repeat homology <3FNG>
F:972-988/Domain: transmembrane #status predicted <TMN>
F:1065-1287/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1282-1304/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1282-1304/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1239/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1245/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.1%; Score 508; DB 1; Length 1337;
Best Local Similarity 38.3%; Pred. No. 1.7e-23;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;

QY 15 QSKKTKKEFAEFLLKQST-----KYKADK-----TYPTVAENAKNKNRKYDILP 65
DB 1017 KSKLRIVENFEAFYKQKQADSCNGFAEYEDLKVGISQPKYAAELAEENRGNRYNNVLP 1076
QY 66 YDYSRVELSLTSDSESSYINANFKYGVKAYIATQGLSTLLDFWRMIWEYSVLII 125
DB 1077 YDYSRVELSLTSDSESSYINANFKYGVKAYIATQGLSTLLDFWRMIWEYSVLII 125
QY 126 VMACMEYEMGKKCYRYAEPGEMOLEFGPFVSVCSEAERKSDYIIRTLKVK--FNSETR 183
DB 1136 IMLTKCVOGRTKCEY--PSKAQDYDITVAMTSEIVLPETVIRDTVKNIQTSESH 1193
QY 184 TIYOHYKKNPDHVPSSIDPILLELWVRCY--QEDDSVPICIHCSACCGRTGVCATV 241
DB 1194 PLRQHFHTSWPDHGVPTDITLINFRLVRYDMKQSPPEPILVHCSAGVGRGTGFIAD 1253
QY 242 DYTWMLLKDCIIPENFVSFLIRENRTQPSLVQTOEQYVLYNNAVLFLFKQMD 296
DB 1254 RLIIQIENENTV----DVYGVYDLMHRPLMVQVEDQYVFLNQCVLDIVRSQKD 1304

RESULT 9
A28334
protein-tyrosine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell variant) - mouse
N:Alternate names: 200K leukocyte common antigen; CD45; PTPY1; T-cell surface glycoprotein; protein-tyrosine-phosphatase (T-cell variant)
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 21-Jan-2000
A:Accession: A28334; A29381; A61180; A60933; A33522; A29075; I54450
R:Thomas, M.L.; Reynolds, P.J.; Chain, A.; Ben-Neriah, Y.; Trowbridge, I.S.
Proc. Natl. Acad. Sci. U.S.A. 84, 5360-5363, 1987
A:Title: B-cell variant of mouse T200 (Ly-5): evidence for alternative mRNA splicing.
A:Reference number: A28334; MUID:87260986
A:Accession: A28334
A:Molecule type: mRNA
A:Residues: 1-1291 <THO>
A:Cross-references: GB:M22455
R:Saga, Y.; Tung, J.S.; Shen, F.W.; Boyse, E.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 6940-6944, 1986
A:Title: Sequences of Ly-5 cDNA: isoform-related diversity of Ly-5 mRNA.
A:Reference number: A29381; MUID:86313686
A:Accession: A29381
A:Molecule type: mRNA
A:Residues: 1-30, 170-517, 'NNT', 521-527, 'G', 529-555, 'S', 557-587, 'S', 589-905, 'Q', 907-93
A:Cross-references: GB:M14342; NID:g198914; PIDN:AAA39458.1; PID:g198915
R:Yi, T.; Cleveland, J.L.; Ihle, J.N.
Blood 78, 2222-2228, 1991
A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells
A:Reference number: A61180; MUID:92032882
A:Accession: A61180
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 730-838 <VIA>
R:Gomez, L.J.; Walker, I.D.; Sandrin, M.S.; McKenzie, I.F.C.
Immunogenetics 25, 263-266, 1987
A:Title: High sequence conservation between rat (T200) and mouse (Ly-5) leukocyte com
A:Reference number: A60933; MUID:87192931
A:Accession: A60933
A:Molecule type: protein
A:Residues: 'R', 289-298, 329, 'V', 331-336, 'Y', 'R', 364-370, 'X', 372-375, 595-608, 638-649, 6
R:Johnson, N.A.; Meyer, C.M.; Pingel, J.T.; Thomas, M.L.
J. Biol. Chem. 264, 6220-6229, 1989
A:Title: Sequence conservation in potential regulatory regions of the mouse and human
A:Reference number: A33522; MUID:89197920
A:Accession: A33522
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <JOH>
A:Cross-references: GB:M22456; NID:g198755; PIDN:AAB46374.1; PID:g554185; GB:J04640;
R:Raschke, W.C.
Proc. Natl. Acad. Sci. U.S.A. 84, 161-165, 1987
A:Title: Cloned murine T200 (Ly-5) cDNA reveals multiple transcripts within B- and T-
A:Reference number: A29075; MUID:87092355
A:Accession: A29075
A:Molecule type: mRNA
A:Residues: 961-1291 <RAS>
A:Cross-references: GB:M15174; NID:g201105; PIDN:AAA40161.1; PID:g201106
R:Tung, J.
Immunogenetics 28, 271-277, 1988
A:Title: Structural features of Ly-5 glycoproteins of the mouse and counterparts in o
A:Reference number: I54450; MUID:88330145
A:Accession: I54450
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 32-73 <RES>
A:Cross-references: GB:M23241; NID:g340850; PIDN:AAA39460.1; PID:g548174
C:Genetics:
A:Gene: Ly-5
C:Superfamily: leukocyte common antigen; leukocyte common antigen cytosolic domain ho
C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-1291/Product: protein-tyrosine-phosphatase (B-cell variant) #status predicted <M
F:24-564/Domain: extracellular #status predicted <EXT>
F:24-30, 170-1291/Product: protein-tyrosine-phosphatase (T-cell variant) #status predi
F:565-586/Domain: transmembrane #status predicted <TMN>
F:587-1291/Domain: intracellular #status predicted <INT>
F:587-1291/Domain: intracellular #status predicted <INT>
F:664-888/Domain: protein-tyrosine-phosphatase homology <PTP>
F:664-888/Domain: protein-tyrosine-phosphatase homology <PTP>
F:664-888/Domain: protein-tyrosine-phosphatase homology <PTP>
F:664-888/Domain: protein-tyrosine-phosphatase homology <PTP>
F:840/Active site: Cys (phosphocysteine intermediate) #status predicted
F:846/Binding site: substrate phosphate (Arg) #status predicted

Query Match 14.0%; Score 507; DB 1; Length 1291;
Best Local Similarity 27.8%; Pred. No. 1.8e-23;
Matches 174; Conservative 84; Mismatches 237; Indels 132; Gaps 24;

QY 1 MDREILQKFLDEAQSCKITKEE--FANFELKLRQSTKYKADTKYPTTVAENAKNKN 58
DB 615 MDVEPIHSDILLETYKRIADGRFLAEFQSIPIRVFSK-----FPIKDKRPHNQKN 668

Qy 282 LVYNAVLEL--FKRQMDVIRDKHSGTESQAKHCEIPEKNHTLQADSYGNLPK-STTKAAK 338
Db 873 LIHQALVEYNQGETEVNLSLHSCLOLKKRDPSPDPLEAE--YQRLPSYRSWRTQH 930
Qy 339 MNQORTKMEIKESSDFRSEISAKBELVLHPAKSSTSDFLNLYSFDKNADT--TM 396
Db 931 IGNOEENKKRRSSNVVPYDFNRVPLKHELEM--SKESEA-----ESDESSDESDSETS 984
Qy 397 KWQTKARPIVCEPLQKHQSLDGLSLLFEGCSNSKP---VNAAGRYFNKSVPIITTKSTPF 453
Db 985 KYINASFW-----SYWPEMMAAQQ-----PIKETIGDFW 1016
Qy 454 ELIQOQRETKEY-----DSKENFSYLESOPHD-----SCFVEMQAQV 490
Db 1017 QMIFORKVKVIMLTMLSGDQVCAQWGEKQTYGDMVYMLKDTNKSAYI-LRAFEL 1075
Qy 491 MHVSSAE-----LNYSLPYDSKHQIRNASNVK 517
Db 1076 RHSKRKEPTVYQYQCTTWKGEELPAEPKDLVTLIQNIK 1114
RESULT 11
JC6312
Protein-tyrosine-phosphatase (EC 3.1.3.48) receptor-type - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jan-2000
C:Accession: JC6312
R:Yang, Y.; Gil, M.C.; Choi, E.Y.; Park, S.H.; Pyun, K.H.; Ha, H.
Gene 186, 77-82, 1997
A:Title: Molecular cloning and chromosomal localization of a human gene homologous to th
A:Reference number: JC6312; MUID:97199372
A:Accession: JC6312
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1440 <YAN>
A:Cross-references: GB:177886
C:Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III repe
-tyrosine-phosphatase homology
C:Keywords: phosphoric monoester hydrolase
F:31-194/Domain: NAM homology <NAM>
F:209-272/Domain: immunoglobulin homology <IMM>
F:294-372/Domain: fibronectin type III repeat homology <3PR>
F:911-1131/Domain: protein-tyrosine-phosphatase homology <PTP>
Query Match 13.68; Score 490.5; DB 2; Length 1440;
Best Local Similarity 27.08; Pred. No. 2.3e-22;
Matches 161; Conservative 91; Mismatches 222; Indels 123; Gaps 22;
Qy 21 KEEFANEFLKLRQSTKYKADKTYPTTVAENAKNIKKRYKDILPYDSYRVELSLITSDE 80
Db 889 KEYESEFF--EGQSASW-----DVAKKDQNRKKNRYGHIIAYDHSRVILQPVEDDP 937
Qy 81 DSSYNANPIKGVYGPKAYIATQGPLSTTLDFWRMIWEYSVLIIVMACMEYEGKKKE 140
Db 938 SSDYINANYIDGYQRPNSYIATQGPVHTVYDFWRMIWEQSACIVMTNLVEVGRVKY 997
Qy 141 RYWAEPGEMOLEFGPFSVCEAEKRKSDYIIR--TLKVKFNSETRTIYQHYKNWPDHDV 198
Db 998 KIWPDDTEV---YGFVKYCTVEMEPALBYVWTFLERRGYNEIREVKQFHTGMPDHGV 1054
Qy 199 PSSIDPILEIWDVRCYQEDDSVPITCIHCSAGCGRTGIVCAIVDYTWMLLKDGIIPEFS 258
Db 1055 PYHATGLLSFIRAVKLSNPPSAGPIVHVCSAGAFRTGCIYIDIMDMAERGVV---D 1110
Qy 259 VFSLIREMRTORPSLVQTOEQVELYNAVLEL-----FK-RQMDVIR----- 299
Db 1111 IYNCVKALRSRINNVQTEQYIFTHDAILEACLGSETAIPVCFEKAAYFDMIDTSQTN 1170
Qy 300 DKHSGTESQAKH-----CIPEKNHTLIQ---ADSYSPN--LPKSTTKAAKMN 341
Db 1171 SSHLKDEFQTLNSVTPRLQAECDSTACLP-RNHDKNRFMDMLPPDRCLPLITIDGESSN 1229

Qy 342 QORTKM--EIKESSDFDRTSEISAKBELVLHPAKSSTSDFLNLYSF----- 388
Db 1230 YINAALMDSYRQPAF-----IVTQYPLPNTVKDFWRLVYDYGCTSTVMLNEV 1277
Qy 389 DKNADTTMKWQTKAF-----PIVGEPLQKHQSLDGLSLLFEGCSNSKPVNA-----A 435
Db 1278 DLSGCGPYWPEEOMLRYPQIVCEMCSMDCDVINRIFRICNITRPQEGYLMVQQFYQL 1337
Qy 436 GRYNSKVPITRTTKSTPFELI-----QORETKVEDSKENFSYLESQPHDSCF-----VE 484
Db 1338 GWASHREVP--GSNSFLKLLIQVEKQOECKEGEGRTIHLNGGSRGNFCAYGIVVE 1395
Qy 485 M-QAQKVMHYSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVENPYFSS 540
Db 1396 MVKRONVV-----DVFHAVKTLRNSKPNNVAEPEQYRFCYDVALEYLES 1439
RESULT 12
S20825
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 6 - human
N:Alternate names: hematopoietic cell phosphatase HCP; protein-tyrosine-phosphatase 1
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 08-Feb-1996 #text_change 11-Jun-1999
C:Accession: B42031; A38189; S20825; S17234; S20837
R:Yi, T.L.; Cleveland, J.L.; Ihle, J.N.
Mol. Cell. Biol. 12, 836-846, 1992
A:Title: Protein tyrosine phosphatase containing SH2 domains: characterization, prefe
A:Reference number: A42031; MUID:92123209
A:Accession: B42031
A:Molecule type: mRNA
A:Residues: 1-595 <YII>
A:Cross-references: GB:M74093
A:Experimental source: T-lymphoid cell line
A:Note: sequence extracted from GenBank
R:Plutsky, J.; Neel, B.G.; Rosenberg, R.D.
Proc. Natl. Acad. Sci. U.S.A. 89, 1123-1127, 1992
A:Title: Isolation of a src homology 2-containing tyrosine phosphatase.
A:Reference number: A38189; MUID:92141214
A:Accession: A38189
A:Molecule type: mRNA
A:Residues: 1-85, 'V', 87-595 <PLU>
A:Cross-references: GB:M77273; NID:g338079; PIDN:AAA36610.1; PID:g338080
A:Note: sequence extracted from NCBI backbone (NCBIN:79619, NCBIP:79620)
R:Shen, S.H.; Bastien, L.; Posner, B.I.; Chretien, P.
Nature 353, 868, 1991
A:Title: Corrigendum: A protein-tyrosine phosphatase with sequence similarity to the
A:Reference number: S20825
A:Accession: S20825
A:Molecule type: mRNA
A:Residues: 1, 'LSRG', 4-595 <SHE>
A:Cross-references: EMBL:X62055; NID:g35781; PIDN:CAA43982.1; PID:g35782
R:Shen, S.H.; Bastien, L.; Posner, B.I.; Chretien, P.
Nature 352, 736-739, 1991
A:Title: A protein-tyrosine phosphatase with sequence similarity to the SH2 domain of
A:Reference number: S17234; MUID:91343005
A:Accession: S17234
A:Molecule type: mRNA
A:Residues: 1, 'LSRG', 4-589, 'VPSRGSERCCPOVAMPQP' <SH2>
A:Experimental source: breast carcinoma cells
A:Note: sequence revised in reference S20805
C:Genetics:
A:Gene: GDB:PTPN6
A:Cross-references: GDB:131389; OMIM:176883
A:Map position: 12p13-12p13
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-pho
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosph
F:4-98/Domain: SH2 homology <SH2A>
F:110-211/Domain: SH2 homology <SH2B>
F:265-521/Domain: phosphatase catalytic domain #status predicted <P>
F:270-504/Domain: protein-tyrosine-phosphatase homology <PTP>
F:453/Active site: Cys (phosphocysteine intermediate) #status predicted
F:459/Binding site: substrate phosphate (Arg) #status predicted

Query Match	13.4%	Score 485;	DB 1;	Length 595;	
Best Local Similarity	33.0%;	Pred. No. 1.4e-22;			
Matches	128;	Conservative 69;	Mismatches 129;	Indels 62;	Gaps
QY	15	OSKKITKEEFANFELKLRKSTQYKADKTYPTTVAENAKNIKNNRYKDLIPDYYSRVLS 74			
DB	235	ESEDTAAGTWEFESLQGEVK-----NLHORLEGORPENKNNRYKNILIPDHSRVILQ 290			
QY	75	LITSD-EDSSYIINANFTKG-VYGP----KAYIAQTGPLSTLLDFWRMIWEXSVLLIWA 128			
DB	291	GRDSNIPGSDYINANYIKNQLLGPDENAKTYIASQGLEATVNDWQMAWQENSRIVMT 350			
QY	129	CMEYEMGKKKCEYBAPPEGMOLFEPFSSCEAEKKSDDYIIRTLKVK--FNSE-TRPI 185			
DB	351	TREVERKNNKCVPIWPEVG-MQRAYGYSVTNGEHDTEYKURTOLOVSPLDNGDLIREI 409			
QY	186	YQEHYKNWPDHVPSSIDPFILELIWDVRCYQED--DSVPICIHCSAGCGRTGVICAIVDY 243			
DB	410	WHYQYLSWPDHGPSEPGGVLSFLDQINQRQESLPHAGPIIIVHCSAGIGRTGTIIVI---466			
QY	244	TWMLLKDGIIPEFNSVPSL-----TREMTRPSLVQTOQEYELVYNVAV---LELF 291			
DB	467	-----DMLMENISTKGLDCDIDIKTIQMVRAQRSGMVQTEAOYKFYVVAIAQFIET 519			
QY	292	KROMVDIRDKHSGETSOAKHICPEKNHTLQADSYSNLPKSTTKAAK-----M 339			
DB	520	KKLEVLQSQ-KQGESEYGN-----TYPAMKNAHAKASRTSSKHKEDVYENL 567			
QY	340	MNOORTKWEIKESSSDFRTSEISAKEE 367			
DB	568	HTNNKREKVKYKQORSADKEKSGSLRRK 595			

RESULT 13
B48148
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type gamma precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: B48148
R:Barnea, G.; Siivonen, O.; Shaanan, B.; Honegger, A.M.; Canolli, P.D.; D'Eustachio, M.L. Cell. Biol. 13, 1497-1506, 1993
A:Title: Identification of a carbonic anhydrase-like domain in the extracellular region
A:Reference number: A48148; MUID:93180796
A:Accession: B48148
A:Status: preliminary
A:Map position: 14
A:Molecule type: mRNA
A:Residues: 1-1442 <BAR>
A:Cross-references: GB:L09562; NID:g293773; PIDN:AAA0022.1; PID:g293774
C:Genetics:
A:Gene: Ptprg
A:Map position: 14
A:Note: extensively polymorphic
C:Superfamily: protein-tyrosine-phosphatase, receptor type gamma; carbonic anhydrase homone-phosphatase homology
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; polymorphism; receptor; transfer F:60-321/Domain: carbonic anhydrase homology <CAH>
F:346-434/Domain: fibronectin type III repeat homology <3FR>
F:871-1105/Domain: protein-tyrosine-phosphatase homology <PTP>
F:1172-1396/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1057/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1063/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match      13.4%; Score 485; DB 1; Length 1442;
Best Local Similarity 25.8%; Pred. No. 5e-22;
Matches 160; Conservative 96; Mismatches 263; Indels 100; Gaps

QY 9 KFLDEAQSKITKEEFANEFKLKROSTYKADKDTPTTVAENAKNIKKRYKDLDPYD 68
      | | | | | : : : : : | | | | | : : : : : | | | | | : :
Db 832 KHIGELYN--SQHGSEDFEEYQRC-----ADMTATGHSNPDNKKNRNRYLILAYDH 885

QY 69 SRVLSLITSDEQ--SSYINANPIKGVGPYKAVIATQGPLSTLLDFWRMIWEYSVLIV 126

```

Db	886	SRVLRPLPGDKSKHSDYINANVVDYGNNAKAYIAATQGPLKSTFFDFWRMWEQNTGII	945
Qy	127	MACMEYEMGKKCKERYWAEPGEMOLEFGPSPVSCYSAEKKSDYIIRTLKV	176
Db	946	MITNLVEGRKKDOYW--PTENTEYGNIIYTLAKTKVHACYTVRRLSVRNTKVKKQK	1003
Qy	177	---KFNSETRITVQFYHKYNPMDHDVPSSIDPFILELIWDVRCYQEDSDVPICHSAGCGR	233
Db	1004	GNPQGRQNETVIQYHYHTQWDPMGVPEYALPVLTVRRSSAARMDDMGFVLVHCSAGVR	1063
Qy	234	TGVICAVDVTWMLLKDGIIIPENFSVFSLIREMARTORPSLVQTOBOYELVYNVAVLELFR	293
Db	1064	TGTY-IVIDSMLOQIKD---KSTVNVVLGELKHIRTQRYLVQTESQYIFIHADALLE---	1115
Qy	294	QMDVIRDKHSGTESQAKH-----CIPEKNHTIQADSDYSPNLPKSTTKAAKMMNQORTKM	347
Db	1116	--AILCKETAVSSSLHSQHSVNSILIPGVGGKTRLEKQFKLITQCNAYVVECFSAQ----	1168
Qy	348	EIKESSDFPRTSEI--SAKEELVLHPAKSSTSFQLELWYSFDKNADTTMKWQTAKPPI	405
Db	1169	--KECNEKRNNSVVAERARVGLAPLPGMKGTDIINASYIMG-----YIRSNEPII	1219
Qy	406	VGEPLQKHQSILDGSLLFECGSSNKPVNAAGRYFNSKVPITTKSTPPELTQORETKVD	465
Db	1220	TQHPL-PHTTKDFWRMTWDH-----NA-----OIVMLPDNQSL	1252
Qy	466	SKENFSYLESOPHSDSCFEVMOAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALG	525
Db	1253	AEDEFVWPSRES--MNCAEATVTLISKDRCLLS-----NEEQIIHDFILEA	1299
Qy	526	VY-SYIPLVENPYFSSWPPSGTSSKMSLDLPKQDGTVPSSLLPTSTSLFSYNSHSS	584
Db	1300	TDDYVLEVRHFQCPKWPNDAPISSTFEL----INVIKEALTRDGTIVHDEYGAVSA	1355
Qy	585	LSLNSPTNIISSLNQESAV	603
Db	1356	GMLCALTTLSQOLENENAV	1374
RESULT 14			
JC5167			
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - chicken			
N:Alternate names: phosphotyrosine phosphatase; PTP1B; PTP2c; SH-PTP2; Syt			
C:Species: Gallus gallus (chicken)			
C:Date: 21-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 16-Jul-1999			
C:Accession: JC5167			
R:Park C.Y.; LaMontagne, K.R.; Tonks, N.K.; Hayman, M.J.			
Gene 177, 93-97, 1996			
A:Title: Cloning and expression of the chicken protein tyrosine phosphatase			
A:Reference number: JC5167; MUID:97080506			

A:Accession: JC5167
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-593 <PAR>
A:Cross-references: GB:U38620; NID:q1054939; PIDN:RAC60049, 1; PID:q1054940
C:Comment: This enzyme plays positive roles in mitogenic signaling and early
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-ty
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specif
E:6-105, 112-193/Domain: SH2 #status predicted <SH2>
E:6-100/Domain: SH2 homology <SH2A>
F:112-210/Domain: SH2 homology <SH2B>
F:273-510/Domain: protein-tyrosine-phosphatase homology <PTP>
E:559-570/Region: proline-rich
E:459/Active site: Cys (phosphocysteine intermediate) #status predicted
F:465/Binding site: substrate phosphate (Arg) #status predicted
F:542, 547-580/Binding site: phosphate(Tyr) (covalent) #status predicted

Query Match	13.4%	Score 483.5;	DB 2;	Length 593;
Best Local Similarity	32.5%;	Pred. No. 1.7e-22;		
Matches 123; Conservative	67;	Mismatches 119;	Indels 69;	Gaps 14;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 12:17:55 ; Search time 39.31 Seconds
(without alignments)
681.606 Million cell updates/sec

Title: US-09-600-358A-4
Perfect score: 3615
Sequence: 1 MDQREILQFLDEAQSKIT.....RPSKSVKLRSPKSGKNFWSL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2356.5	65.2	802	1 PTPN_MOUSE	P29352 mus musculus
2	1157	32.0	780	1 PTPN_MOUSE	Q05209 homo sapien
3	1130.5	31.3	775	1 PTPN_MOUSE	P35831 mus musculus
4	516.5	14.3	1304	1 CD45_HUMAN	P08575 homo sapien
5	513	14.2	1238	1 PTPJ_MOUSE	Q84455 mus musculus
6	508	14.1	1337	1 PTPJ_MOUSE	Q12913 homo sapien
7	499	13.8	1152	1 CD45_MOUSE	P06800 mus musculus
8	493	13.6	1255	1 CD45_MOUSE	P04157 rattus norv
9	492.5	13.6	1439	1 PTPK_HUMAN	Q15262 homo sapien
10	485	13.4	595	1 PTPG_MOUSE	P29350 homo sapien
11	485	13.4	1442	1 PTPG_MOUSE	Q05909 mus musculus
12	479	13.3	1452	1 PTPM_MOUSE	P28828 mus musculus
13	479	13.3	2316	1 PTP2_MOUSE	Q62656 rattus norv
14	478.5	13.2	1457	1 PTPK_MOUSE	P35822 mus musculus
15	478	13.2	2314	1 PTPM_HUMAN	P23471 homo sapien
16	477	13.2	1452	1 PTPM_HUMAN	P28827 homo sapien
17	476	13.2	1445	1 PTPG_HUMAN	P23470 homo sapien
18	473.5	13.1	593	1 PTPN_MOUSE	P41499 rattus norv
19	473	13.1	593	1 PTPN_MOUSE	Q06124 homo sapien
20	471	13.0	1912	1 PTPD_HUMAN	P23468 homo sapien
21	468	12.9	595	1 PTPM_MOUSE	P29351 mus musculus
22	467.5	12.9	1897	1 PTPF_HUMAN	P10586 homo sapien
23	466.5	12.9	585	1 PTPN_MOUSE	P35235 mus musculus
24	463.5	12.8	521	1 PTPD_MOUSE	P34137 dictyostell
25	462.5	12.8	1301	1 PTPN_MOUSE	P35832 drosophila
26	460	12.7	802	1 PTPN_MOUSE	P18433 homo sapien
27	460	12.7	829	1 PTPN_MOUSE	P18052 mus musculus
28	456	12.6	1422	1 PTPG_CHICK	Q08936 gallus gall
29	452	12.5	1705	1 PTPG_MOUSE	P70289 mus musculus
30	446.5	12.4	1462	1 PTP6_MOUSE	P16620 drosophila
31	441	12.2	796	1 PTPN_MOUSE	Q03348 rattus norv
32	439.5	12.2	2029	1 LAR_MOUSE	P16621 drosophila
33	436.5	12.1	434	1 PTPN_MOUSE	Q13016 gallus gall

34	436	12.1	979	1 PTPN_MOUSE	Q06073 mus musculus
35	435	12.0	359	1 PTPN_MOUSE	P49445 rattus norv
36	434	12.0	1630	1 PTPN_MOUSE	P35992 drosophila
37	433	12.0	979	1 PTPN_MOUSE	P36722 bos taurus
38	433	12.0	983	1 PTPN_MOUSE	Q03259 rattus norv
39	432	12.0	979	1 PTPN_MOUSE	Q16849 homo sapien
40	430.5	11.9	432	1 PTPN_MOUSE	P20417 rattus norv
41	428	11.8	1711	1 PTPN_MOUSE	Q04612 rattus norv
42	427.5	11.8	435	1 PTPN_MOUSE	P18031 homo sapien
43	427.5	11.8	913	1 PTPN_MOUSE	P26045 homo sapien
44	427	11.8	926	1 PTPN_MOUSE	P29074 homo sapien
45	426.5	11.8	360	1 PTPN_MOUSE	P35236 homo sapien

ALIGNMENTS

RESULT	1
PTN8_MOUSE	
ID	PTN8_MOUSE
AC	P29352; STANDARD; PRT; 802 AA.
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-DEC-1992 (Rel. 24, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP (EC 3.1.3.48).
GN	PTP8
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	MEDLINE=92235615; PubMed=1373816;
RT	Mathews R.J., Bowne D.B., Flores E., Thomas M.L.;
RT	"Characterization of hematopoietic intracellular protein tyrosine phosphatases: description of a phosphatase containing an SH2 domain and another enriched in proline-, glutamic acid-, serine-, and threonine-rich sequences";
RT	Mol. Cell. Biol. 12:2396-2405(1992).
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein tyrosine + phosphate.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-1- TISSUE SPECIFICITY: SPLEEN, THYMUS, LYMPH NODE, AND BONE MARROW.
CC	-1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL: M90388; AAA39994.1; .
DR	PUR: B44390; B44390.
DR	HSP: Q06124; 2SHP.
DR	MGD: MGI:107170; Ptpn8.
DR	InterPro: IPR000387; Tyr_phosphatase.
DR	InterPro: IPR000242; Tyr_prot_phptase.
DR	Pfam: PF00102; Y_phosphatase; 1.
DR	PRINTS: PR00700; PRTYPHPTASE.
DR	SMART: SM00194; PTPc; 1.
DR	PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
FT	DOMAIN 23 288 PROTEIN-TYROSINE PHOSPHATASE.
FT	ACT_SITE 227 227 BY SIMILARITY.
SQ	SEQUENCE 802 AA; 89714 MW; 0FIE45339BD4613E CRC64;

Query Match

65.2%; Score 2356.5; DB 1; Length 802;

```
Best Local Similarity 68.5%; Pred. No. 1.5e-137;
Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

QY 1 MDQREILQKFLDEAQSCKITKEEFANEFLLKFKQSTKYKADKTYPTTVAENAKNIKNRY 60
DB 1 MDQREILQQLLKQAKKKUNSEFSEFLKFKQSTKYKADKTYPTTVAQRPNKIKNNRY 60
QY 61 KDILPYDSRVLSLTSDESDSYINANFIKGYGPKAYIATQGPLSTTLTLLDFWRMIWEY 120
DB 61 KDILPYDHSVLSSLTSDESDSYINAFIKGYGPKAYIATQGPLSTTLTLLDFWRMIWEY 120
QY 121 SVLIIVMACMEYEMGKKCRYWAEPEGEMOLEFGPFSVSCAEAKRKS DYIIRTLKVKFNS 180
DB 121 RILVIVMACMEFEMGKKCRYWAEPEGETQLQFGPFSISCEAEKKS DYKIRTLKAKFNN 180
QY 181 ETRTIQFHYKNPDHDPVSSIDPILLEIWDVRCYQEDDSVPICIHCSACGRTGVICAI 240
DB 181 ETRIIQFHYKNPDHDPVSSIDPILQIWDVRCYQEDDCVPICIHCSACGRTGVICA - 239
QY 241 VDTWMLLKDGIIPEFNSVFLSIREMTQRPVSLVQTOEQYELVYNNAVLEFKRQMDVIRD 300
DB 240 VDTWMLLKDGIIPEFNSVFLSIREMTQRPVSLVQTOEQYELVYNNAVLEFKRQMDVIRD 299
QY 301 KHGTSQAKHCIEPKNHTLOADSYSNLPKSTTKAAMNQ---QRTKMEIKESSSDF 357
DB 300 NHLGREIQACCSIQEQLSTVEADSCPLDLPKNAMRDYKTTNQHSGQGAESTGSSGL 359
QY 358 RTSISAKELVLHPAKSSYFDFLELNYSFQDNADTKMWTQKAPPIVGEPLQKHQSLD 417
DB 360 RTSTMNAREELVLHSAKSSFNCLNCCNKKAVITRNGQARSPVGEPLQKHQSLD 419
QY 418 LGSLLFFGCSNKPVNAGRYFNKVPITRTKSTPPELIQOQRETKYDSENFSYLESQP 477
DB 420 FGSMLFGCSPALPINTADRYHNSGPKVKTSTPPELIQOQRTNDLAVGDFGSCLESQ 479
QY 478 HDSCFV-EMQAQKVMHVSSELNLSYLDKSKHQIRNASNVKHHDSALGVYSIPLVNP 536
DB 480 HEHYSRLRELQVQVYHVSSEELNLSL- ----GACDASCVPSPHSGALRVHLTSLAEDP 534
QY 537 YFSSWPPSGTSSKMSLDLPEKQDQTVPPSSLLPTSSYFNSSLSLSPNTNSSL 596
DB 535 YFSSPPNSADSKMSFDLPEKQDQATSGALLPASSITSFYFNPHDSLNTLTSFPP 594
QY 597 LNOESAVLATAPRIDDEIPPLPVRTPEFIVVEAGEFSPNPKSLSSAVKVKIGTSLE 656
DB 595 LNQETAVEAPSRDDEIPPLPTEFIVVEAGEFSPNPKSLSSAVKVKIGTSLE 652
QY 657 WGTSEPKKFDSDSVILRPSKSVKLRSPKSGKN 688
DB 653 CSGTSEMKS-HDSVGFTPSKNVLRSPKSDRH 683

RESULT 2
PTNC_HUMAN STANDARD; PRT; 780 AA.
AC Q05209; Q16130;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein-tyrosine phosphatase, non-receptor type 12 (EC 3.1.3.48)
DE (Protein-tyrosine phosphatase G1) (PTPG1).
GN PTPN12.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=93112015; PubMed=1472029;
RA Takekawa M., Itoh F., Hinoda Y., Arimura Y., Toyota M., Sekiya M.,
RA Adachi M., Imai K., Yachi A.;
RA "Cloning and characterization of a human cDNA encoding a novel
```



```
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILDLWVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFQNESRLYQHYVWVNDHVPSSIDPILDLWVRCYQEDDSVPICIHCSAGCGRTGA 240
QY 237 ICAVDYTWMLLKDGILIPENSVFSLIREMRTQPSLQTOEQYELVYNAVLFRQMD 296
Db 241 ICAI-DYTNWLLKAGKIPPEFNVLQEMRTQPSLQTOEQYELVYNAVLFRQMD 299
QY 297 V-----IRD--KHSQTESQAKHCIPKKNHTLQADSYSPNLPKSTTKAAKMNQRTK 346
Db 300 LYIEHGAQKIADGVNEINTENMVSSIEPEKQ-----DSSPPKPPRTS-----CLVEGD 350
QY 347 MEIKE-----SSSDFRSTSEISAKEEL-----VLHPAKSSTSFDELELNTS 387
Db 351 EEILQPPPEHPVPPILTPSPSAFTVTTWODNDRYHHPKPVHLVMSSEQHSADLNRNYS 410
QY 388 -----FDKNADTTMKQTKAFPIVGEPLQKHOSLDGLSLLFEG----- 425
Db 411 KSTELPGKNESTIEQDKKLERNLSPFIKVPLOQEP-----KSFQDNTLLNRGHAIKIKS 466
QY 426 ---C---SNSKP-----VNAAGRYNSKVPITRTKS-----TPFELIQOERTK----- 462
Db 467 ASPIADKISKQOELSSDNLVGDTSNCSVDCSVTOSNVKSVVTPPEESONSOTPPRPDL 526
QY 463 EVDKSNFYLESQPSHDSFCFVEMQAKVMHVSSAEALNY-----SLPYDSKKHQIRNASNV 516
Db 527 PLDEKGHVTSFHGPENA---IPIPLSEGNSSDINYQTRKTVSLTPSPTTQVETPDLV 582
QY 517 KHDSSALGVYIPLVNPENYSSWPPSGTSSKMSLDLPEK-ODGVYFVSSLLPTSTSL 575
Db 583 DHNTS-----PLFRPLSFTNPLHSDDS-----DSDERNSDGAVTOKNTNISTASAT 630
QY 576 FSYNHSLSLSPNTNISLLNOESAVLATAPR-----IDDETTPPLPVRTPESTFVVEE 631
Db 631 VSAATSEISIRKVLPMISIRHNIAGTHSGAEKVDVSEDSPPPLPERTPEFVLAEE 690
QY 632 AGEFSPNVKSLSSAVKVKIGTSLWEG-----GTSEPKFDDSVILRPSKSVKLRSKPSG 686
Db 691 -----HNTPE-----VRSESELQSQERSQKKSEG---LITSENEKCDHPAGS 730
QY 687 KNF 689
Db 731 IHY 733

RESULT 3
PTNC_MOUSE STANDARD; PRT; 775 AA.
AC P35831;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Protein-tyrosine phosphatase P19 (EC 3.1.3.48) (P19-PTP) (MPTP-PEST).
GN PTPN12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92272714; PubMed=1590786;
RA den Hertog J., Pals C.E., Jonk L.J., Kruijer W.;
RT "Differential expression of a novel murine non-receptor protein
RT tyrosine phosphatase during differentiation of P19 embryonal
RT carcinoma cells.";
RL Biochem. Biophys. Res. Commun. 184:1241-1249(1992).
RN [2]
RP REVISIONS TO 297-416.
RX MEDLINE=93112015; PubMed=1472029;
RA Takekawa M., Itoh F., Hinoda Y., Arimura Y., Toyota M., Sekiya M.,
RA Adachi M., Imai K., Yachi A.;
RT "Cloning and characterization of a human cDNA encoding a novel
RT putative cytoplasmic protein-tyrosine-phosphatase.";
```

```
RL Biochem. Biophys. Res. Commun. 189:1223-1230(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=95289971; PubMed=7772023;
RA Charest A., Wagner J., Shen S.H., Tremblay M.L.;
RT "Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein
RT tyrosine phosphatase.";
RL Biochem. J. 308:425-432(1995).
CC -I- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
CC tyrosine + phosphate.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X63440; CAA45037.1; ALT_SEQ.
DR EMBL; X86781; CAA60477.1; -.
DR HSSP; Q06124; 2SHP.
DR MGD; MGI:104673; Ptpn12.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPHTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT DOMAIN 27 291 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 231 231 BY SIMILARITY.
FT CONFLICT 296 296 K -> N (IN REF. 1).
FT CONFLICT 328 332 KODSP -> DETS (IN REF. 1).
FT CONFLICT 380 380 W -> V (IN REF. 1).
SQ SEQUENCE 775 AA; 86992 MW; 7106D73F5014E411 CRC64;

Query Match 31.3%; Score 1130.5; DB 1; Length 775;
Best Local Similarity 36.6%; Pred. No. 3.4e-62;
Matches 280; Conservative 110; Mismatches 165; Indels 209; Gaps 28;

QY 1 MDQREILQKFLDEAQSCKIT----KEEFANFLKLRQSTKYKADKTYPTTVAENAKNIK 56
Db 1 MEQVEILRRFIQIRVOAMKSPDHNGEDNFARDFMRLRLSTKYRTEKIYPTATGEENVK 60
QY 57 KNRYKDILPDYSRVLSITDESDSYINANFIKGVYGPAYATQGPLSTLLDFWRM 116
Db 61 KNRYKDILPDHDSRVKLTLPKTSQSDSYINANFIKGVYGPAYATQGPLSTLLDFWRM 120
QY 117 IWEYSVLIVMACMEYEMGKCKERYWAEFGEMQLEFGPFSYSCAEKKRSDYIIITLKV 176
Db 121 IWEYNVMIVMACREFEMGKCKERYWPLYGDEPTIFAFPKISCENEQARTFIETLL 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILDLWVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFQNESRLYQHYVWVNDHVPSSIDPILDLWVRCYQEDDSVPICIHCSAGCGRTGA 240
QY 237 ICAVDYTWMLLKDGILIPENSVFSLIREMRTQPSLQTOEQYELVYNAVLFRQMD 296
Db 241 ICAI-DYTNWLLKAGKIPPEFNVLQEMRTQPSLQTOEQYELVYNAVLFRQMD 299
QY 297 V-----IRDKHSQTESQAKHCIPKKNHTLQADSYSPNLPKSTTKAAKMNQRTK 348
Db 300 LYIEHGAQKIADGVNEINTENMVSSIDSE----KQDSSPPKPPRTS-----CLVEGDAKEE 351
QY 349 IKE-----SSSDFRST-----EISAKELVLHPAKSSTSF 379
```


Db 666 PIKARPPFNQNKRYVDILPYDYNRVLSEINGDAGSNYNASYIDGFKPRKYIAAQ 725
QY 105 PLSTLLDFWRMIWEYSVLIIIVACMEYEMGKKCEYWAEPCEMQLFPGFVSCEAEK 164
Db 726 PRDETVDFFWRMIWEQATVYVWTRCEEGRNKKACIYWSMEGTRAFGDVGVVVKINQHK 785
QY 165 RKSDYIIRTLKV---KFNSETRTIYQPHYKNWPDHVPSSIDPILILEIWDVRCVQEDDSV 221
Db 786 RCPDYIIQKLVNKKKATGREVTHIQFTSWPDHGVPEDPHLLKLLRRVNAFNSFSG 845
QY 222 PICHSAGCGRGVCAIVDYVWMLKDGIIIPEN-PSVFSLIEMRTQPSLVQTOEQY 280
Db 846 PIVVHCSSAGVGTGTIGI-----DAMLEGLEAENKVDYVGYVVKLRQRCLMVQVEAQY 900
QY 281 ELVYNVAVLEL--FKRQMDVIRDKHSGTESQAKHCIPKNTKLOADSYSPNLPK-STTKAA 337
Db 901 ILIHQALVEYNQGETEVNSELHPYLHNKKRDPSPSEPLEAE--FQRLPSYRSWRTQ 958
QY 338 KMNQORTMEIKESSFDFRTSEISAKELVL-----HPAKSSTSDFFLELNYSFDKNA 392
Db 959 HIGNOEENKSNRNSVPIPDYNRVPLKHELENSKEHSDSDSDSDSDSESPSKYINA 1018
QY 393 DTMK-WQTAFFIVGEPLOKHOSLDGLSLF-----EGCSNKPVN 433
Db 1019 SFINSYWKPEVMIAAQGL-KETIGDFWQIFQKVKVIVMLTELKHGDOEICAQ---YW 1074
QY 434 AAGR--YFNSKVPITRT-KSTPEL--IQORETKVDNSKENFSY-----LESOPHD 479
Db 1075 GEGQTYGDIEVDLKDFTKSTTLRVFELHRSKRDSRTVYQYQYNNWSVEQLPAEPKE 1134
QY 480 -SCFVEMQAQVHVHSAELNYSIPYDSKHQIRNASNVKHHDSALGVY 527
Db 1135 LISMIQVVKQKPKNSSEGN-----KHKSTPLLIH 1166

RESULT 5
PTPJ_MOUSE STANDARD; PRT; 1238 AA.
ID PTPJ_MOUSE
AC O64455;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Protein-tyrosine phosphatase eta precursor (EC 3.1.3.48) (R-PTP-eta)
DE (HPTP beta-like tyrosine phosphatase).
GN PTPRJ OR BYP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRL-LPR/LPR;
RX MEDLINE=96140699; PubMed=8549806;
RA Kuramochi S., Matsuda S., Matsuda Y., Saitoh T., Ohsugi M.,
Yamamoto T.;
RT "Molecular cloning and characterization of Byp, a murine
receptor-type tyrosine phosphatase similar to human DEP-1";
RL FEBS Lett. 378:7-14(1996).
CC -/- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
tyrosine + phosphate.
CC -/- SUBCELLULAR LOCATION: Type I membrane protein.
CC -/- TISSUE SPECIFICITY: EXPRESSED IN EVERY TISSUE EXAMINED.
CC -/- SIMILARITY: CONTAINS 1 PROTEIN-TYROSINE PHOSPHATASE DOMAIN.
CC -/- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D45212; BAA08146.1; -.
DR HSSP; P18052; 1YFO.
DR GSD; MGI:104574; Ptpnj.
DR InterPro; IPR003961; Ptpnj.
DR InterPro; IPR000387; Tyr_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PTPPHPTASE.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00506; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00505; TYR_PHOSPHATASE_PTP; 1.
KW Signal; Glycoprotein; Transmembrane; Repeat; Hydrolase.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 1238 PROTEIN-TYROSINE PHOSPHATASE ETA.
FT DOMAIN 29 876 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 877 897 POTENTIAL.
FT DOMAIN 898 1238 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 40 112 FIBRONECTIN TYPE-III 1.
FT DOMAIN 121 260 FIBRONECTIN TYPE-III 2.
FT DOMAIN 268 348 FIBRONECTIN TYPE-III 3.
FT DOMAIN 356 434 FIBRONECTIN TYPE-III 4.
FT DOMAIN 442 518 FIBRONECTIN TYPE-III 5.
FT DOMAIN 529 608 FIBRONECTIN TYPE-III 6.
FT DOMAIN 966 1238 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 1140 1140 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 506 506 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 538 538 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 572 572 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 662 662 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 668 668 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 691 691 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 838 838 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1238 AA; 136782 MW; 939479EDC8016835 CRC64;

Query Match 14.2%; Score 513; DB 1; Length 1238;
Best Local Similarity 39.0%; Smad, No. 5.9e-24;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

QY 15 QSKKITKEEFANFLKLRQST-----KYADK---TYPTVAENAKNIKKNRYDILP 65

Qy 126 VMACMEYEMGRKKCERYWAEPCEMOLEFSPVSCAEARKSDYIIRTLKVK--FNSETR 183
 Db 1136 IMLTKCVQGRKCEYW--FSKQADYGDITVAMTSEIVLPETWIRDTFTYKNTQTSSEH 1193
 Qy 184 TIYQFHYKNWPDHVPSSIDILELWDRY--QEDSVFICHCAGCGRTGVICAI 241
 Db 1194 PLRQFHFTSWPDHGVDPDTLLINFRYLVRDYMKSPPSPILVHCSAGVGRGTFTIAID 1253
 Qy 242 DYTMLLKDGILIPENFVSFLIREMRTORPSILVOTQOYELVYNVAVLELFRQMD 296
 Db 1254 RLIIQIENENTV----DVIGIVIDLRMRPLMVOTEDQYVFLNOCVLDIVRSQKD 1304

RESULT 7
 CD45_MOUSE
 ID CD45_MOUSE STANDARD; PRT; 1152 AA.
 AC P06800;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte common antigen precursor (EC 3.1.3.48) (L-CA) (Lymphocyte
 DE common antigen Ly-5) (CD45) (T200).
 GN PTPRC OR Ly-5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86313368; PubMed=2944116;
 RA Saga Y., Tung J.-S., Shen F.-W., Boyse E.A.;
 RA "Sequences of Ly-5 cDNA: isoform-related diversity of Ly-5 mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6940-6944(1986).
 RN [2]
 RP REVISIONS.
 RA Saga Y., Tung J.-S., Shen F.-W., Boyse E.A.;
 RA Proc. Natl. Acad. Sci. U.S.A. 84:1991-1991(1987).
 RN [3]
 RP SEQUENCE OF 10-124 FROM N.A.
 RC TISSUE-T-cell;
 RX MEDLINE=86042665; PubMed=3864163;
 RA Shen F.-W., Saga Y., Litman G., Freeman G., Tung J.-S., Cantor H.,
 RA Boyse E.A.;
 RT "Cloning of Ly-5 cDNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7360-7363(1985).
 RN [4]
 RP SEQUENCE OF 822-1152 FROM N.A.
 RX MEDLINE=87092355; PubMed=2948186;
 RA Raschke W.C.;
 RT "Cloned murine T200 (Ly-5) cDNA reveals multiple transcripts within
 B- and T-lymphocyte lineages";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:161-165(1987).
 CC -1- RECEPTOR. REQUIRED FOR T-CELLS ACTIVATION THROUGH THE ANTIGEN
 CC RECEPTOR. THE FIRST PTPASE DOMAIN HAS ENZYMIC ACTIVITY, WHILE
 CC THE SECOND ONE SEEMS TO AFFECT THE SUBSTRATE SPECIFICITY OF THE
 CC FIRST ONE.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF ISOFORMS ARE PRODUCED BY
 CC ALTERNATIVE SPLICING.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS RESTRICTED TO THE HEMATOPOIETIC
 CC COMPARTMENT OF DEVELOPMENT.
 CC -1- PTM: HEAVILY N- AND O-GLYCOSYLATED.
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; M14342; AAA39458.1; --
 DR EMBL; M11934; AAA39461.1; --
 DR EMBL; M15174; AAA40161.1; --
 DR PIR; A29381; A29381.
 DR HSP; P18052; 1YFO.
 DR MGD; MGI:97810; Ptprc.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR000387; Tyr_phosphatase.
 DR InterPro; IPR000242; Tyr_prot_phptase.
 DR Pfam; PF00041; fn3; 3.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00700; PRTYPHPTASE..
 DR SMART; SM00194; PTEC; 2.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
 KW Glycoprotein; Transmembrane; Phosphorylation; B-cell; T-cell; Repeat;
 KW Alternative splicing; Hydrolase; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 1152 LEUCOCYTE COMMON ANTIGEN.
 FT DOMAIN 24 425 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 426 447 POTENTIAL.
 FT DOMAIN 448 1152 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 233 329 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 330 421 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 520 769 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 811 1084 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 701 701 BY SIMILARITY.
 FT ACT_SITE 1016 1016 BY SIMILARITY.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1152 AA; B4D956B4E32EA812 CRC64;

Query Match 13.8%; Score 499; DB 1; Length 1152;
 Best Local Similarity 27.6%; Pred. No. 3.8e-23;
 Matches 173; Conservative 83; Mismatches 239; Indels 132; Gaps 24;

Qy 1 MDQREILQFLDEAQSCKITKEE--FANEFLKLRQSTKYKADKTYPTTVAENAKTKKN 58
 Db 476 MDVEPIHSDILETYVKRKIADEGRFLAEFQSIPIRVFSK-----FPIKDARKPHNQKN 529
 Qy 59 RYKDILPYDYSRVELSLTSDSDSSYNANFIKGYGPKAVIATQGPLSTLLDFWRMIW 118
 Db 530 RYVDILPYDYNRVELSEINGDAGSTYINASYIDGFKPERKYIAAAGRPDEYVDFWRMIW 589
 Qy 119 EYSVLITVMACMEVEMGKKCKERYWAEPCEMOLEFSPVSCAEARKSDYIIRTLKVKF 178
 Db 590 EQKATVIVWVTRCEGGRNKKCAEYWPSEGTAFKDIVVTINDHKRCPDYIIQKLVAVH 649
 Qy 179 NSET---RTIYQFHYKNWPDHVPSSIDIPLELWDRYQEDSVFICHCAGCGRTG 235
 Db 650 KKERATGREVTHIQTSWPDHGVDPEDPHLLKLRNRNNAFNSFFSGPIVHCSAGVGRGT 709
 Qy 236 VICAIYDVTWMLLKDGILPE-NFSVFLIREMRTQPSLQVOTQOYELVYNVAVLEL--FK 292
 Db 710 TYIGI-----DAMLEGLEAEGKVDVGYVVKLRQRCIMQVQVEAQYILIHQALVEYNQFG 764
 Qy 293 RQMDVIRDKHSGTSSQAKHCIPENKHTLOADSYSPNLPK-STTKAAMNQORTKMEIKE 351


```

Db      565  EQQLVDRDEEKQLINVDPIHSDLLLETYYKRIADGRFLTAEFOSIDPRVFESK-----F 618
Qy      45   PTTVAENAKNIKNRYKDILPDYDSRVLSLTISDESDSYINANIKGVYGPKAYIATQG 104
Db      619  PKDARKSQNQNRYVDILPDYNRVSELSEINGDAGSTYINASYIDGKEPRKYIAAQG 678
Qy      105  PLSTTLDFWRMTWEYSVLIIYVACHMEYEMGKKKCKERTWAEPEGMOLEGFEGFSPVSCRAEK 164
Db      679  PROETVDVDFWKMIWEQKATVIVMVTFRCEBGNRNKCAEYMPCEMEGTRTFRDVVVYTIIDHK 738
Qy      165  RKSDYIIRTLKVFNSET--RTIYQFYHKWNPDDHDVSSDIPLELILWDVRCYQEDDSV 221
Db      739  RCPDYIIQKLSIAHKKEKATGREVTHIQTISWPDHGVDPEDPHLLKLKRRRNAFNSFFSG 798
Qy      222  PICHSAGCGRTGVCAIVDYTWMLLKDGIIIPENFSVFSILIREMRTQPSLIVQTQEQYE 281
Db      799  PIVVHCSAGVGRGTGYIGIDAMLESLEAGKV---DYVGYVNLRRORCLMVQVEAQYI 854
Qy      282  LVYNVAVLEL--FKRQMDVIRDKHSGETSQAQKICIEPKNHTLQAOADSYSNLPK--STTKAAK 338
Db      855  LIHQALVEYNQFQETEVNLSLSELHSQLNKKRPDPSPLEAE--YQRLPSYRSWRQTQ 912
Qy      339  MNQORTKMEIKESSDFPRTSEISAKELVLHPAKSSTSDFDFLELNVSPFKNADT--TM 396
Db      913  IGNOEENKKNRSNNVYPDFNRVPLKHELEM--SKESEA---ESDESDDEDSDSETS 966
Qy      397  KWOTKAPPIVGEPLQHQSLDLGLSLLFEGCSNSKP---VNAAGRYFNSKVPITRTKSTPF 453
Db      967  KYNASFVM-----SWKPEMMIAAAG-----PLKETIGDFW 998
Qy      454  ELIQOQRETKEV-----DSKENFSYLESQPHD-----SCFVEMQAQKV 490
Db      999  QMIFQKRVKVIYVMLTSLMSDQOEVCQAQYWGECQKQYGDMEVWLKDTNKKSSAYI-LRAFEL 1057
Qy      491  MHYSSAP-----LNTSLPYDSKHQIRNASNVK 517
Db      1058  RHSKRKEPRTVYQYQCTTWKGELPAEPKDLVTLIQNIK 1096

RESULT      9
PTPK_HUMAN  STANDARD;          PRT; 1439 AA.
AC      Q15262; Q14763;
ID      DT 01-NOV-1997 (Rel. 35, Created)
DT      DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Protein-tyrosine phosphatase kappa precursor (EC 3.1.3.48) (R-PTP-
DE      kappa).
OS      PTPRK OR PTPK.
GN      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96279245; PubMed=8663237;
RA      Fuchs M., Mueller T., Lerch M., Ullrich A.;
RT      "Association of human protein-tyrosine phosphatase kappa with members
RT      of the armadillo family.";
RL      J. Biol. Chem. 271:16712-16719(1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Foreskin;
RX      MEDLINE=97199372; PubMed=9047348;
RA      Yang Y., Gil M.C., Choi E.Y., Park S.H., Pyun K.H., Ha H.;
RT      "Molecular cloning and chromosomal localization of a human gene
RT      homologous to the murine R-PTP-kappa, a receptor-type protein
RT      tyrosine phosphatase.";
RL      Gene 186:77-82(1997).
CC      -1- FUNCTION: REGULATION OF PROCESSES INVOLVING CELL CONTACT AND
CC      ADHESION SUCH AS GROWTH CONTROL, TUMOR INVASION, AND METASTASIS.
CC      FORMS COMPLEXES WITH BETA-CATENIN AND GAMMA-CATENIN/PLAKGLOBIN.

```

```
FT CARBOHYD 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 552 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 586 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 9 L -> V (IN REF. 2).
FT CONFLICT 158 S -> T (IN REF. 2).
FT CONFLICT 284 A -> P (IN REF. 2).
FT CONFLICT 422 T -> S (IN REF. 2).
FT CONFLICT 672 AEL -> CRT (IN REF. 2).
FT CONFLICT 715 S -> T (IN REF. 2).
FT CONFLICT 732 A -> AA (IN REF. 2).
FT CONFLICT 1366 E -> K (IN REF. 2).
SQ SEQUENCE 1439 AA; 162087 MW; EED529AF7C9F4451 CRC64;

Query Match 13.6%; Score 492.5; DB 1; Length 1439;
Best Local Similarity 29.1%; Pred. NO. 1.3e-22;
Matches 136; Conservative 77; Mismatches 167; Indels 87; Gaps 16;

QY 21 KEFANEFLKRSQTKYKADYTTVAENAKNKKRYKDILPDYRSVELSLITSD 80
Db 888 KEYESFF---EGQSASW-----DVAKKDQNRKAKRYGNIAYDHSRVLQPVDDP 936
QY 81 DSSYINANFIKGVGKAYIATQGPLSTLLDFWMIWEYSVLIIWACMEYEMGKKCE 140
Db 937 SSDYINANYIDGYQRPVSHYIATQGPVHEVDFWMIWEQSACIVMTNLVEGRVKCY 996
QY 141 RYWAEPGEQMLEFGPFSVCEAEKRSYIIR--TLKVKFNSETRIYQHYKNPDRHV 198
Db 997 KYWPDDEV---YGFVKVTCVEMELAERYVVRTFLRRGYNEIREVKQFHTGWPDHGV 1053
QY 199 PSSIDPILEIWDVRCYQEDDSVPICHSACCGRTGVCIAIVDTWMLLKDGIIPEFNS 258
Db 1054 PYHATGLLSFIRRLKSLNPPSAGPIVHVSAGAGRTGCVIYIDIMDMAERGVV---D 1109
QY 259 VESLREMRTRQPSLVQTOEQVELYVNAVLEL-----FK-RQMDVIR----- 299
Db 1110 IYNCVKALRSRINNVQVEEQYIFTHDAILEACLGCTAIPVCFKAAFYDMIRDSQTN 1169
QY 300 DKHSGTESQAKH-----CIPEKNHTIQ--ADSVSPN--LPKSTTKAAKMMN 341
Db 1170 SSHLKDEFOTLNSVTPRLQAEDCSIACLP-RNHDKNRMDMLPPDKRCLPLFTIDGESSN 1228
QY 342 QORTKM--EIKESSFDRTSISAKEELVHPAKSSTSFDFLELNSF----- 388
Db 1229 YTNAALMDSYRQPAAF-----IVTQYPLPNTVKDFWRLVYDYGCTSIYMLNEV 1276
QY 389 DKNADTTMKWQTKAF-----PIVGEPLQKHQSLDGLSLLFEGCSNSKP 431
Db 1277 DLSQSCPQWPEGMLRGPIQVECMSCMDCDVINRIFRICNLTRP 1323

RESULT 10
ID PTN6 HUMAN
AC P29350; STANDARD; PRT; 595 AA.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase, non-receptor type 6 (EC 3.1.3.48)
DE (Protein-tyrosine phosphatase 1C) (PTP-1C) (Hematopoietic cell
DE protein-tyrosine phosphatase) (SH-PTP1) (Protein-tyrosine phosphatase
DE SHP-1).
GN PTPN6 OR PTPIC OR HCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
```

```
RX MEDLINE=92123209; PubMed=1732748;
RA Yi T., Cleveland J.L., Ihle J.N.;
RT "Protein tyrosine phosphatase containing SH2 domains:
RT characterization, preferential expression in hematopoietic cells, and
RT localization to human chromosome 12p12-pl3.";
RL Mol. Cell. Biol. 12:836-846(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE-Breast;
RX MEDLINE=91343005; PubMed=1652101;
RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
RT "A protein-tyrosine phosphatase with sequence similarity to the SH2
RT domain of the protein-tyrosine kinases.";
RL Nature 352:736-739(1991).
RN [3]
RP REVISIONS.
RA Shen S.H., Bastien L., Posner B.I., Chretien P.;
RL Nature 353:868-868(1991).
RN [4]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE=92141214; PubMed=1736296;
RA Plutzky J., Neel B.G., Rosenberg R.D.;
RT "Isolation of a src homology 2-containing tyrosine phosphatase.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1123-1127(1992).
RN [5]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RX MEDLINE=95394454; PubMed=7665165;
RA Banville D., Stocco R., Shen S.H.;
RT "Human protein tyrosine phosphatase 1C (PTPN6) gene structure:
RT alternate promoter usage and exon skipping generate multiple
RT transcripts.";
RL Genomics 27:165-173(1995).
RN [6]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RX MEDLINE=96303695; PubMed=8723724;
RA Ansari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E.,
RA Spanos S., Malley T., Gibbs R.A.;
RT "A gene-rich cluster between the CD4 and triosephosphate isomerase
RT genes at human chromosome 12p13.";
RL Genome Res. 6:314-326(1996).
RN [7]
RP PHOSPHORYLATION.
RX MEDLINE=95300784; PubMed=7781604;
RA Li R.Y., Gails F., Ragab A., Ragab-Thomas J.M.F., Chap H.;
RT "Tyrosine phosphorylation of an SH2-containing protein tyrosine
RT phosphatase is coupled to platelet thrombin receptor via a pertussis
RT toxin-sensitive heterotrimeric G-protein.";
RL EMBO J. 14:2519-2526(1995).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 248-399.
RX MEDLINE=98447672; PubMed=9774441;
RA Yang J., Liang X., Niu T., Meng W., Zhao Z., Zhou G.W.;
RT "Crystal structure of the catalytic domain of protein-tyrosine
RT phosphatase SHP-1.";
RL J. Biol. Chem. 273:28199-28207(1998).
CC -|- FUNCTION: PLAYS A KEY ROLE IN HEMATOPOIESIS. THIS PTPASE ACTIVITY
CC MAY DIRECTLY LINK GROWTH FACTOR RECEPTORS AND OTHER SIGNALING
CC PROTEINS THROUGH PROTEIN-TYROSINE PHOSPHORYLATION. THE SH2 REGIONS
CC MAY INTERACT WITH OTHER CELLULAR COMPONENTS TO MODULATE ITS OWN
CC PHOSPHATASE ACTIVITY AGAINST INTERACTING SUBSTRATES.
CC -|- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O ~ protein
CC tyrosine + phosphate.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE)
CC AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -|- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEMATOPOIETIC
CC CELLS.
CC -|- PTM: PHOSPHORYLATED ON SERINE AND TYROSINE.
CC -|- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
CC -|- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
-----
```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M74903; AAA35963.1; -
DR EMBL: X62055; CAA43982.1; -
DR EMBL: M77273; AAA36610.1; -
DR EMBL: U15528; AAA82880.1; -
DR EMBL: U15536; AAA82880.1; JOINED.
DR EMBL: U15535; AAA82880.1; JOINED.
DR EMBL: U15534; AAA82880.1; JOINED.
DR EMBL: U15533; AAA82880.1; JOINED.
DR EMBL: U15532; AAA82880.1; JOINED.
DR EMBL: U15531; AAA82880.1; JOINED.
DR EMBL: U15530; AAA82880.1; JOINED.
DR EMBL: U15529; AAA82880.1; JOINED.
DR EMBL: U15528; AAA82879.1; -
DR EMBL: U15537; AAA82879.1; JOINED.
DR EMBL: U15535; AAA82879.1; JOINED.
DR EMBL: U15534; AAA82879.1; JOINED.
DR EMBL: U15533; AAA82879.1; JOINED.
DR EMBL: U15532; AAA82879.1; JOINED.
DR EMBL: U15531; AAA82879.1; JOINED.
DR EMBL: U15530; AAA82879.1; JOINED.
DR EMBL: U15529; AAA82879.1; JOINED.
DR EMBL: U47924; AAB51323.1; -
DR EMBL: U47924; AAB51322.1; -
DR PIR: S20825; S20825.
DR PIR: B42031; B42031.
DR PDB: 1GWZ; 22-AUG-99.
DR PDB: 1FPR; 07-MAR-01.
DR MIM: 176883; -
DR InterPro: IPR000980; SH2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00102; Y_phosphatase; 1.
DR PRINTS: PR00700; PTYPHPTASE.
DR PRINTS: PR00401; SH2DOMAIN.
DR SMART: SM00194; PTPC; 1.
DR SMART: SM00252; SH2; 2.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 1.
DR PROSITE: PS50001; SH2; 2.
KW Hydrolase; SH2 domain; Repeat; Phosphorylation; Alternative splicing;
3D-structure. 4 100 SH2 1.
FT DOMAIN 110 213 SH2 2.
FT DOMAIN 269 514
FT ACT_SITE 453 453
FT VARSPLOC 1 39
FT VARSPLOC 40 44
FT CONFLICT 2 3
FT CONFLICT 6 86
FT CONFLICT 86 146
FT CONFLICT 146 146
SQ SEQUENCE 595 AA; 67561 MW; 4D7736C21D3542D2 CRC64;
MISSING (IN SHORT ISOFORM).
SLSVR -> MLRSG (IN SHORT ISOFORM).
VR -> LSRG (IN REF. 2).
H -> L (IN REF. 5).
L -> V (IN REF. 4).
V -> E (IN REF. 5).

Query Match 13.4%; Score 485; DB 1; Length 595;
Best Local Similarity 33.0%; Pred. No. 1.1e-22;
Matches 128; Conservative 69; Mismatches 129; Indels 62; Gaps 14;
QY 15 QSKKTKKEFANEFKLKRSQSKYKADTKYPTVAENAKNKNRYKDLIPYDYSRVELS 74
Db 235 ESDTAKAGFWNEEFSLQKQYK-----NLHORLEGORPENKGNRYKNILPFDHSRVLIQ 290
QY 75 LETSD-EDSSYINANFIKG-VYGP-----KAYIATQGLSTLLDFWRMIWEYSVLIIVMA 128

Db 291 GRDSNIPGSDYINANYIKNQLGPDENAKTYIASQCLEATVNDWFQMAQWNSRVIVMT 350
QY 129 CMYEMGKKKCRRYNAEPGEQMLEFGPFSVCEAKRKSDDYIIRLKVK--FNSE--TRTI 185
Db 351 TREVEKGRNKCVPYVPEVG-MQRAYGPYSVTNCGEHDTTTEYKLRITLQVSPDLNDLIREI 409
QY 186 YGFHYKNPDDHVPSSIDPILILEIWDVRCYQED--DSVPICIHCSAGCGRTGVICAIVDY 243
Db 410 WHYQLSWPDHGVSEPGGVLSFLQDINORQESLPHAGPIIVHCSAGIGRTGIIVI--- 466
QY 244 TWMLLKDGIIPEFNSVFSL-----IREMRTORPSLVQTEQYELVYNAV---LELF 291
Db 467 -----DMLMENISTKGLDCDDIDIOKTIQMVRAQRSGMVQTEAQYKFIYVAIAQFIET 519
QY 292 KRQMDVIRDKHSGTESQAKHCIPKKNHITLOADSYSPNLKPKSTTKAAK-----M 339
Db 520 KKKLEVLSQ-KGQSEYGN-----TYPPAMKNAHAKASRTSSKHEDVYENL 567
QY 340 MNQORTKMEIKESSSDFDPTSEISAKEE 367
Db 568 HTYKNREKVKQKRSADREKSKGSLKRK 595
RESULT 11
ID PTPG_MOUSE STANDARD; PRT: 1442 AA.
AC Q05909;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE protein-tyrosine phosphatase gamma precursor (EC 3.1.3.48) (R-PTP-gamma).
DE gamma).
GN PTPRG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93180796; PubMed=8382771;
RA Barnea G., Silvennoinen O., Shaanan B., Honegger A.M., Canoll P.D.,
RA D'Eustachio P., Morse B., Levy J.B., Laforgia S., Huebner K.,
RA Muscatello J.M., Sap J., Schlessinger J.;
RT "Identification of a carbonic anhydrase-like domain in the
RT extracellular region of RPTP gamma defines a new subfamily of receptor
RT tyrosine phosphatases";
RL Mol. Cell. Biol. 13:1497-1506(1993).
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: DETECTED IN BRAIN, LUNG, KIDNEY, HEART, LIVER,
CC SKELETAL MUSCLE, SPLEEN AND TESTES. IT IS DEVELOPMENTALLY
CC REGULATED IN THE BRAIN.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE EUKARYOTIC-
CC TYPE CARBONIC ANHYDRASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF00194; carb_anhydrase; 1.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00102; Y_phosphatase; 2.
DR PRINTS: PR00700; PRTYPHPHTASE.
DR ProDom: PD000865; Carb_anhydrase; 1.
DR SMART: SM00060; FN3; 1.
DR SMART: SM00194; PTPC; 2.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
DR PROSITE: PS00055; TYR_PHOSPHATASE_PTP; 2.
KW Glycoprotein; Transmembrane; Hydrolase; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 1442
FT DOMAIN 20 733
FT TRANSMEM 734 759
FT DOMAIN 760 1442
FT DOMAIN 56 322
FT DOMAIN 347 441
FT DOMAIN 866 1122
FT DOMAIN 1123 1442
FT ACT_SITE 1057 1057
FT SITE 1348 1348
FT CARBOHYD 109 109
FT CARBOHYD 113 113
FT CARBOHYD 156 156
FT CARBOHYD 359 359
FT CARBOHYD 444 444
FT CARBOHYD 719 719
FT SEQUENCE 1442 AA; 161242 MW; 5887715568BECDB CRC64;

Query Match 13.4%; Score 485; DB 1; Length 1442;
Best Local Similarity 25.8%; Pred. No. 3.8e-22;
Matches 160; Conservative 96; Mismatches 263; Indels 100; Gaps 19;

QY 9 KFLDEAQQKTKTEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKKNRYKDILPYD 68
DB 832 KHIGELYSN--SQGFSEDFEEVQRCF-----ADNNTAHSNHPDNKKNRYINLAYDH 885
QY 69 SRVSLSDTDED--SSYINANFTKGVGPKAYATATOGPLSTLLDFWRMIWEYSVLIV 126
DB 886 SRVLRPLPGKDSKHSYDIINANYVDGYNKAKYATATOGPLKSTPDEFWRMIWEQNTGIII 945
QY 127 MACHEYEMGKKCRYNAEPCGEMOLGPFVSCEAEKRKSDYIIRTKV-----176
DB 946 MITNLVEKGRKKCDQYW--PTENTEENIIVLUKSTKVHACYTVRLSVNRTKVKKGOK 1003
QY 177 ---KFNSETRRTIYOFHYKNWPDHDPSSIDPILSLIWDVRCYQEDDSVPICIHCSAGCR 233
DB 1004 GNPGRQNERVTIYHYTQWDMGPEYALPVLTFVRRSSAARMPDMPGLVHCSAGVR 1063
QY 234 TGVTCIVDYTMLLKQIGITPENSFVSLIREMTQRPISVQTOEQYELVYNVAVLEFKR 293
DB 1064 TGTY-IVIDSMLOQIKD---KSTVNLGFLKHIRTQRYNLVQTEEQYIFIHIDALLE---- 1115
QY 294 QMDVIRDKHSCTESQAKH-----CIPEKHTLQADSVSPNLKSTTKAAKMMNQORTKM 347
DB 1116 ---AILKETAVSSQLSHSYNSILIPGVGCKTRLEQFKLITOCNAKYVBCFSAQ---- 1168
QY 348 EIKESSSDFRTSEI--SAKEELVHLPAKSTSFDFLELANSFQKNADTMMKWTKAFPI 405
DB 1169 --KECNKERNSSVPAERARVGLAPLPGMKGTDIYNASYIMG-----YYRSNEFII 1219
QY 406 VGEPLQKHQSLDLSGLLFECCGSKSPVNAAGRYFNSKVPITRTKSTPELLQORETKEVD 465
DB 1220 TQHPPL-PHTTKDFRMIMDWH-----NA-----QIIVMLPDNQSL 1252
QY 466 SKENFYSLESPHDSFCFEMOAKVMHVSSAELNYSLPYDSKHOIRNASNVKHHDSALG 525
DB 1253 AEDEFVYWPSSREES---MNCEAFVTILSKDRCLIS-----NEEQIIIHDFILEA 1299
```

```
QY 526 VY-SYIPLVENPYFSSWPPSGTSSKMSLDLPEKQDGTVPFSSLLPTSSSTSLFSSYNHSS 584
DB 1300 TODDYVLEVRHFQCPKPNPDAPISSTFEL-----INVKEALTRDGPPTIVHDEYGAWS 1355
QY 585 LSLNSPTNISSLLNQESAV 603
DB 1356 GMLCAUUTLSQLENAV 1374

RESULT 12
PTPM_MOUSE
ID PTPM_MOUSE STANDARD; PRT; 1452 AA.
AC P28828;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein-tyrosine phosphatase MU precursor (EC 3.1.3.48) (R-PTP-MU).
GN PTPRM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RC MEDLINE=92008644; PubMed=1655529;
RA Gebbink M.F.B.G., van Etten I., Hateboer G., Suijkerbuijk R.,
RA Beijersbergen R., Geurts van Kessel A., Moolenaar W.H.;
RT "Cloning, expression and chromosomal localization of a new putative
RT receptor-like protein tyrosine phosphatase.";
RL FEBS Lett. 290:123-130(1991).
CC -!- FUNCTION: MAY PLAY A KEY ROLE IN SIGNAL TRANSDUCTION AND GROWTH
CC CONTROL.
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN LUNG, LESS IN BRAIN AND
CC HEART.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: X58287; CAA41225.1; -.
DB PIR: S17670; S17670.
DB HSP; P28827; IRPM.
DB MGD; MGI:102694; Pcprrm.
DB InterPro: IPR003961; FN_III.
DB InterPro: IPR003962; FNIII_repeat.
DB InterPro: IPR003599; Ig.
DB InterPro: IPR003006; Ig_MHC.
DB InterPro: IPR000998; MAM.
DB InterPro: IPR000387; TYR_phosphatase.
DB InterPro: IPR000242; Tyr_prot_phptase.
DB Pfam: PF00041; fn3; 3.
DB Pfam: PF00047; Ig; 1.
DB Pfam: PF00629; MAM; 1.
DB Pfam: PF00102; Y_phosphatase; 2.
DB PRINTS: PR00014; FNTYPEIII.
DB PRINTS: PR00020; MAMDOMAIN.
DB PRINTS: PR00700; PRTYPHPHTASE.
DB SMART: SM00060; FN3; 2.
DB SMART: SM00409; IG; 1.
DB SMART: SM00137; MAM; 1.
DB SMART: SM00194; PTPC; 2.
```

DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE: PS00055; TYR_PHOSPHATASE_2; 2.
 DR PROSITE: PS00740; MAN_1; 1.
 DR PROSITE: PS00060; MAN_2; 1.
 KW Hydrolase; Receptor; Glycoprotein; Signal; Transmembrane;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 20
 FT CHAIN 21 1452
 FT DOMAIN 21 742
 FT TRANSMEM 743 764
 FT DOMAIN 765 1452
 FT DOMAIN 22 184
 FT DOMAIN 199 267
 FT DOMAIN 287 374
 FT DOMAIN 383 466
 FT DOMAIN 486 571
 FT DOMAIN 589 671
 FT DOMAIN 923 1153
 FT DOMAIN 1213 1447
 FT ACT_SITE 1095 1095
 FT ACT_SITE 1389 1389
 FT DISULFID 206 260
 FT CARBOHYD 72 72
 FT CARBOHYD 92 92
 FT CARBOHYD 131 131
 FT CARBOHYD 249 249
 FT CARBOHYD 406 406
 FT CARBOHYD 414 414
 FT CARBOHYD 454 454
 FT CARBOHYD 534 534
 FT CARBOHYD 544 544
 FT CARBOHYD 598 598
 FT CARBOHYD 651 651
 FT CARBOHYD 681 681
 SQ SEQUENCE 1452 AA; 163594 MW; F99D0FC67922CFIE CRC64;
 Query Match 13.3%; Score 479; DB 1; Length 1452;
 Best Local Similarity 35.4%; Pred. No. 9e-22; 98; Indels 20; Gaps 4;
 Matches 96; Conservative 57; Mismatches 98;
 QY 21 KEAFNEFLKLRSTKTKYKADKTYPTTVAENAKNKKRYKIDLPYDYSRVLSLITDSE 80
 Db 901 KEYESFF-----EGSAPWDSAKDENMKRNGNIAYDSRVQLMEGDN 949
 QY 81 DSSYNANFIVGPKYATQGPLSTLLDFWRMIWEYSVLLIIVMACMEYEMGKKKE 140
 Db 950 NSDYINGNYIDGHRPNHYIATQGPMTQETIYDFWRMWHENTASIIIMVNLVEGVRKCC 1009
 QY 141 RYWAEPGEWQLEFGFPFSYSCAEKRSYIIRTLKVKNS--ETRTIYQFHYKNPDDHV 198
 Db 1010 KYWPDTEI---YKDIKVTLDITELLAEVYVTFVAVKRGHIEIRIQFHTGWPDDHV 1066
 QY 199 ESSIDPILIEWRCYQEDDSVPICICHSACGRTGVCIVDVTMLLKDGIIPEFNS 258
 Db 1067 PYHATGLGCFVQVQKSPNAGPLVHCSAGAGTGCFTVIDIMDMAERGVV---D 1122
 QY 259 VPSLTREMTQPSLVQTOEQVELYNAVLE 289
 Db 1123 IYNCVRELSRRVNMVQTEQYVFHDAILE 1153
 RESULT 13
 PTP2_RAT
 ID PTP2_RAT
 AC Q62656; Q62621; STANDARD; PRT; 2316 AA.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase zeta precursor (EC 3.1.3.48) (R-PTP-
 DE zeta) (Phosphacan) (3F8 chondroitin sulfate proteoglycan)
 DE sulfate proteoglycan).

GN PTPRZ1 OR PTPRZ OR PTPZ.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT FORMS).
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Brain;
 RX MEDLINE-96063026; PubMed-7579589;
 RA Maurel P., Meyer-Puttlitz B., Flad M., Margolis R.U.,
 RT "Nucleotide sequence and molecular variants of rat receptor-type
 RL protein tyrosine phosphatase-zeta/beta.";
 RN DNA Seq. 5:323-328(1995).
 [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (PHOSPHACAN).
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Brain;
 RX MEDLINE-94195772; PubMed-7511813;
 RA Maurel P., Rauch U., Flad M., Margolis R.K., Margolis R.U.;
 RT "Phosphacan, a chondroitin sulfate proteoglycan of brain that
 RL interacts with neurons and neural cell-adhesion molecules, is an
 RT extracellular variant of a receptor-type protein tyrosine
 RT phosphatase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2512-2516(1994).
 [3]
 RP BINDING TO N-CAM AND NG-CAM (PHOSPHACAN).
 RC MEDLINE-95096181; PubMed-7528221;
 RA Milev P., Friedlander D.R., Sakurai T., Karthikeyan L., Flad M.,
 RT Margolis R.K., Grumet M., Margolis R.U.;
 RT "Interactions of the chondroitin sulfate proteoglycan phosphacan, the
 RL intracellular domain of a receptor-type protein tyrosine phosphatase,
 RL with neurons, glia, and neural cell adhesion molecules.";
 RN J. Cell Biol. 127:1703-1715(1994).
 [4]
 RP BINDING TO TENASCIN (PHOSPHACAN).
 RC MEDLINE-94216329; PubMed-7512960;
 RA Grumet M., Milev P., Sakurai T., Karthikeyan L., Bourdon M.,
 RT Margolis R.K., Margolis R.U.;
 RT "Interactions with tenascin and differential effects on cell adhesion
 RL of neurocan and phosphacan, two major chondroitin sulfate
 RL proteoglycans of nervous tissue.";
 RL J. Biol. Chem. 269:12142-12146(1994).
 [5]
 RP INTERACTION WITH CONTACTIN.
 RC MEDLINE-95354206; PubMed-7628014;
 RA Peles E., Nativ M., Campbell P.L., Sakurai T., Martinez R., Lev S.,
 RA Clary D.O., Schilling J., Barnea G., Plowman G.D., Grumet M.,
 RA Schlesinger J.;
 RT "The carbonic anhydrase domain of receptor tyrosine phosphatase beta
 RL is a functional ligand for the axonal cell recognition molecule
 RL contactin.";
 RL Cell 82:251-260(1995).
 CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SPECIFIC
 CC DEVELOPMENTAL PROCESSES IN THE CNS.
 CC -!- FUNCTION: PHOSPHACAN, PREVIOUSLY DESIGNATED 3F8 CHONDROITIN
 CC SULFATE PROTEOGLYCAN OR 3H1 KERATAN SULFATE PROTEOGLYCAN DEPENDING
 CC ON THE GLYCOSYLATION STATUS, IS A SOLUBLE NERVOUS TISSUE-SPECIFIC
 CC PROTEOGLYCAN. IT IS SYNTHESIZED BY GLIA AND BINDS TO NEURONS AND
 CC BUT NOT TO LAMININ AND FIBRONECTIN. PHOSPHACAN ACTS AS A POTENT
 CC INHIBITOR OF CELL ADHESION AND NEURITE OUTGROWTH.
 CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -!- SUBUNIT: THE CARBONIC-ANHYDRASE LIKE DOMAIN BINDS TO CONTACTIN.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: A LONG FORM (SHOWN HERE), A SHORT FORM
 CC MISSING AN INTERNAL SEGMENT, AND PHOSPHACAN WHICH REPRESENTS THE
 CC ENTIRE EXTRACELLULAR DOMAIN ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: NERVOUS TISSUE-SPECIFIC.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE EUKARYOTIC-
 CC TYPE CARBONIC ANHYDRASE FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

Db 2164 LTIQDFILEATQDDYVLEV--RHFQCPKWPDPSPISKT-----FELI 2204

Search completed: May 26, 2002, 12:26:53
Job time: 538 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 12:16:55 ; Search time 95.55 Seconds
(without alignments)
1252.878 Million cell updates/sec

Title: US-09-600-358A-4
Perfect score: 3615
Sequence: 1 MDQREILQKFLDEAQSCKIT.....RPSKSVKLRSKGNFWSL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3611	99.9	692	4	O95064
2	3569	98.7	808	4	O95063
3	3512.5	97.2	807	4	O9Y2R2
4	2971.5	82.2	799	4	O9P0U2
5	1042	28.8	382	11	O63745
6	748	20.7	453	11	P70602
7	735.5	20.3	453	11	O922E3
8	733.5	20.3	453	11	O61152
9	708	19.6	458	4	O99952
10	515.5	14.3	1114	4	O9H0V6
11	515.5	14.3	1143	4	Q16614
12	513	14.2	361	11	O61373
13	511	14.1	1195	13	O9PUM2
14	508	14.1	1216	11	O62884
15	508	14.1	1337	4	Q15255
16	507	14.0	1291	11	Q61812

17	498	13.8	492	4	O75664
18	498	13.8	1435	11	O9JIZ2
19	498	13.8	1445	11	O99M81
20	498	13.8	1454	11	O99M82
21	498	13.8	1463	4	O43655
22	494	13.7	1444	11	O9JIZ1
23	492	13.6	1444	11	O9JLP0
24	490	13.6	1216	13	O9IBD8
25	488.5	13.5	2271	13	O9I909
26	487.5	13.5	832	13	O90947
27	485	13.4	597	4	O969V8
28	483.5	13.4	593	13	O90687
29	483.5	13.4	1200	13	O9I054
30	482	13.3	624	4	O9UK67
31	480.5	13.3	487	5	O9NL13
32	479	13.3	1455	11	O99M80
33	478	13.2	1648	4	O9UDR6
34	477.5	13.2	1788	13	O9IAJ0
35	477	13.2	613	11	P81718
36	475.5	13.2	508	13	O90VJ5
37	475.5	13.2	1237	13	O91976
38	474.5	13.1	468	13	O91BA5
39	474.5	13.1	1115	4	O9HD43
40	473	13.1	470	5	O9NL15
41	473	13.1	1406	13	O9W6V5
42	472.5	13.1	438	11	O64642
43	471.5	13.0	597	11	O64509
44	471	13.0	1214	5	O9VAL3
45	470.5	13.0	468	13	O91BA2

ALIGNMENTS

RESULT 1

O95064 PRELIMINARY; PRT; 692 AA.
AC O95064;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LYPHOID PHOSPHATASE LYP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Roifman C.M.;
RT "cDNA of Lyp2 Protein Tyrosine Phosphatase.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001847; AAD00905.1; -
DR HSSP; Q06124; 2SHP.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_prot_phptase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00194; PTPc; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_ptp; 1.
KW Hydrolase.
SQ SEQUENCE 692 AA; 78828 MW; 543ED06AE9A8C4DF CRC64;

Query Match 99.9%; Score 3611; DB 4; Length 692;

Best Local Similarity 99.9%; Pred. No. 5.9e-249;

Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDQREILQKFLDEAQSCKITKEEFANEFKLKRQSTYKADKYPTTVAENAKIKNNRY 60

Db 1 MDQREILQKFLDEAQSCKITKEEFANEFKLKRQSTYKADKYPTTVAENAKIKNNRY 60

QY 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
Db 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGPFSVCEAEKRKSDYIIRTLKVKENS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGPFSVCEAEKRKSDYIIRTLKVKENS 180
QY 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
Db 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
QY 241 VDYTWMLLKDGIIIPENFSVSLIREMRTQPSLVQEQYELVYNAVLELFRQMDVIRD 300
Db 241 VDYTWMLLKDGIIIPENFSVSLIREMRTQPSLVQEQYELVYNAVLELFRQMDVIRD 300
QY 301 KHSGETSAKHCIPEKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
Db 301 KHSGETSAKHCIPEKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
QY 361 EISAKEELVLPKAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSLDLGS 420
Db 361 EISAKEELVLPKAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSLDLGS 420
QY 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
Db 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
QY 481 CFVEMQAOKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPIVLPENPFSS 540
Db 481 CFVEMQAOKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPIVLPENPFSS 540
QY 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
Db 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
QY 601 SAVLATAPRIDDEIPPLPVPTPESFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
Db 601 SAVLATAPRIDDEIPPLPVPTPESFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
QY 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
Db 661 SEPCKFDDSVILRPSKSVKLRSPKS 685

RESULT 2
Q95063 PRELIMINARY; PRT; 808 AA.
ID O95063
AC O95063;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE LYMPHOID PHOSPHATASE LypI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Roifman C.M.;
RT "Human cDNA of LypI protein Tyrosine Phosphatase.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF001846; AAD00904.1; --
DR HSP: Q06124; 2SHP.
DR InterPro: IPR000387; TYR_phosphatase.
DR InPro: IPR000242; Tyr_prot_phptase.
DR Pfam: PF0102; Y_phosphatase; 1.
DR PRINTS: PR00700; PRTPHPTASE.
DR SMART: SM00194; PTPC; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.

SQ SEQUENCE 808 AA; 91763 MW; 8375848DF8D3D853 CRC64;
Query Match 98.7%; Score 3569; DB 4; Length 808;
Best Local Similarity 99.9%; Pred. No. 7.le-246;
Matches 684; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDQREILQKDFDEAQSKKITKEEFANEFKLKRSQTKYKADKTYPTTVAENAKNIKKRY 60
Db 1 MDQREILQKDFDEAQSKKITKEEFANEFKLKRSQTKYKADKTYPTTVAENAKNIKKRY 60
QY 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
Db 61 KDILPYDSRVLSLITSDSSSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
QY 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGPFSVCEAEKRKSDYIIRTLKVKENS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGPFSVCEAEKRKSDYIIRTLKVKENS 180
QY 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
Db 181 ETRTIYQHYKNWPDHVPSSIDPILWVRCYQEDDSVPICIHCSAGCGRTGVCICAI 240
QY 241 VDYTWMLLKDGIIIPENFSVSLIREMRTQPSLVQEQYELVYNAVLELFRQMDVIRD 300
Db 241 VDYTWMLLKDGIIIPENFSVSLIREMRTQPSLVQEQYELVYNAVLELFRQMDVIRD 300
QY 301 KHSGETSAKHCIPEKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
Db 301 KHSGETSAKHCIPEKNHTLQADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFDRTS 360
QY 361 EISAKEELVLPKAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSLDLGS 420
Db 361 EISAKEELVLPKAKSTSFDFLELNYSPDKNADTTMKWTKAFPIVGPLOKHQSLDLGS 420
QY 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
Db 421 LFFEGCSNSKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDKSNFYLESQPHDS 480
QY 481 CFVEMQAOKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPIVLPENPFSS 540
Db 481 CFVEMQAOKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYIPIVLPENPFSS 540
QY 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
Db 541 WPPSGTSSKMSLDLPEKODGTVPFSSLLPTSTSLFSYNSHSSLSLNSPTNISLLNQE 600
QY 601 SAVLATAPRIDDEIPPLPVPTPESFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
Db 601 SAVLATAPRIDDEIPPLPVPTPESFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWGGT 660
QY 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
Db 661 SEPCKFDDSVILRPSKSVKLRSPKS 685
RESULT 3
Q9Y2R2 PRELIMINARY; PRT; 807 AA.
ID Q9Y2R2
AC Q9Y2R2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE HOMOLOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu T., Zhang J., Fu G., Zhang Q., Ye M., Zhou J., Wu J., Shen Y.,
RA Yu M., Chen S., Mao M., Chen Z.;
RT "Human protein tyrosine phosphatase (70zpep) homolog.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF077031; AAD2764.1; -

DR HSP: A06124; 2SHP.

DR InterPro: IPR000387; TYR_phosphatase.

DR Pfam: PF00102; Y_phosphatase; 1.

DR PRINTS: PR00700; PRTVPHPTASE.

DR SMART: SM00194; PTPC; 1.

DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.

DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.

KW Hydrolase.

SQ SEQUENCE 807 AA; 91708 MW; A5B4B4CAB795E812 CRC64;

Query Match 97.2%; Score 3512.5; DB 4; Length 807;

Best Local Similarity 98.8%; Pred. No. 7.5e-242;

Matches 677; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 MDQREILQKFLDEAQSCKITKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60

DB 1 MDQREILQKFLDEAQSCKITKEEFANFLKLRQSTKYKADKTYPTTVAEKPNKIKKNRY 60

QY 61 KDILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLDFWMIWEY 120

DB 61 KDILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLDFWMIWEY 120

QY 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGFSPVSCAEKRSKSDYIIRTLVKVFN 180

DB 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGFSPVSCAEKRSKSDYIIRTLVKVFN 180

QY 181 ERTTIYQHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240

DB 181 ERTTIYQHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240

QY 241 VDTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLFRQMDVIRD 300

DB 241 -DYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLFRQMDVIRD 299

QY 301 KHSGETSQAKHCIPKKNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 360

DB 301 KHSGETSQAKHCIPKKNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 359

QY 361 ELSAKEELVHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 420

DB 361 ELSAKEELVHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 419

QY 421 LFFEGCSNKPVNAAGRYFNKVPITRTKSTPFELIQORETKEVDSENFYSLESOPHDS 480

DB 421 LFFEGCSNKPVNAAGRYFNKVPITRTKSTPFELIQORETKEVDSENFYSLESOPHDS 479

QY 481 CFVEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 540

DB 481 CFVEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 539

QY 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNNHSSLSLNSPTNISLLNOE 600

DB 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNNHSSLSLNSPTNISLLNOE 599

QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 660

DB 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 659

QY 661 SEPKFDDSVILRPSKVKLRSPKS 685

DB 661 SEPKFDDSVILRPSKVKLRSPKS 684

RESULT 4

Q9P002

ID Q9P002

AC Q9P002;

DT 01-OCT-2000 (Tremblrel. 15, Created)

PRELIMINARY; PRT; 799 AA.

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE PROTEIN TYROSINE PHOSPHATASE.

GN PTP.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBL_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=ADRENAL GLAND;

RA Huang C., Zhang C., Wu T., Peng Y., Gu Y., Zhang L., Jiang C., Li Y.,

RA Han Z., Wang Y., Chen Z., Fu G.;

RT "A novel gene expressed in the human adrenal gland";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR HSP: A06124; 2SHP.

DR InterPro: IPR000387; TYR_phosphatase.

DR PRINTS: PR00700; PRTVPHPTASE.

DR SMART: SM00194; PTPC; 1.

DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.

DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.

KW Hydrolase.

SQ SEQUENCE 799 AA; 90610 MW; 93F5385016F33D0C CRC64;

Query Match 82.2%; Score 2971.5; DB 4; Length 799;

Best Local Similarity 87.3%; Pred. No. 2.6e-203;

Matches 598; Conservative 17; Mismatches 61; Indels 9; Gaps 9;

QY 1 MDQREILQKFLDEAQSCKITKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60

DB 1 MDQREILQKFLDEAQSCKITKEEFANFLKLRQSTKYKADKTYPTTVAEKPNKIKKNRY 60

QY 61 KDILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLDFWMIWEY 120

DB 61 KDILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLDFWMIWEY 120

QY 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGFSPVSCAEKRSKSDYIIRTLVKVFN 180

DB 121 SVLIIVMACMEYEMGKKKERYWAEPMQLEFGFSPVSCAEKRSKSDYIIRTLVKVFN 180

QY 181 ERTTIYQHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240

DB 181 ERTTIYQHYKKNPDHVPSSIDPILLELWDVRCQEDDSVPICHCAGCGRTGVICAI 240

QY 241 VDTYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLFRQMDVIRD 300

DB 241 -DYTMLLKDGIIIPENFVSFLIREMRTQPSLVQTQEQYELVYNAVLFRQMDVIRD 299

QY 301 KHSGETSQAKHCIPKKNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 360

DB 301 KHSGETSQAKHCIPKKNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSDFRTS 359

QY 361 ELSAKEELVHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 420

DB 361 ELSAKEELVHPAKSSTSFDFLELNYSPDKNADTTMKWTKAFPIVGEPLQKHQSLDLS 419

QY 421 LFFEGCSNKPVNAAGRYFNKVPITRTKSTPFELIQORETKEVDSENFYSLESOPHDS 480

DB 421 LFFEGCSNKPVNAAGRYFNKVPITRTKSTPFELIQORETKEVDSENFYSLESOPHDS 471

QY 481 CFVEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 540

DB 481 CFVEMQAQKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS 531

QY 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNNHSSLSLNSPTNISLLNOE 600

DB 541 WPPSGTSSKMSLDLPEKQDGVFPSSLLPTSTSLFSYNNHSSLSLNSPTNISLLNOE 591

QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKICTSLEWGT 660

Db 592 SAVLATAPRIDDEIPLHFLXGTSPSFVWEAGEFSPNPKSLSSAVKVGKIGTSLWGGT 651
Qy 661 SEPKKFDSDVILRPSKSVKLRSPKS 685
Db 652 SEPKKFDSDVILRPSKSVKLRSPKS 676
RESULT .5
Q63745
ID Q63745 PRELIMINARY; PRT; 382 AA.
AC Q63745;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=95046282; PubMed=7957881;
RA Moriyama T., Kawanishi S., Inoue T., Inai E., Kaneko T., Xia C.,
RA Takenaka M., Kanada T., Ueda N.;
RT "cDNA cloning of a cytosolic protein tyrosine phosphatase (RKPTP) from
RT rat kidney.";
RL FEBS Lett. 353:305-308(1994).
DR EMBL; D38072; BAA07266.1; -.
DR HSP; Q06124; 2SHP.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTPHPHTASE.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 382 AA; 44438 MW; 25A517FDEBB7C344 CRC64;

Query Match 28.8%; Score 1042; DB 11; Length 382;
Best Local Similarity 53.5%; Pred. No. 2.8e-66;
Matches 201; Conservative 62; Mismatches 89; Indels 24; Gaps 5;
Qy 1 MDQREILQKFLDERAQSKKIT----KEEFANFLKLRQSTKYKADKTYPTVAENAKNIK 56
Db 1 MEQVEILRRFQRYQAMKSPDHNGEDNFARDFMLRLRLSTKYRTEKIYPTATGEKENVK 60
Qy 57 KNRVKDILPYDYSRVLSLITSDSDSSYINANFTKGVYGPAYATATOGPLSTTLTDFWRM 116
Db 61 KNRKDLILPFDHRSRVKLLTKTPSQSDYINANFTGVYGPAYATOGPLANTVIDFWRM 120
Qy 117 IWEXSVLIIVMACMEYEMGKKKERYNAEPGEMQLEFGPFSVSCAEAKRKS DYIIRTLKV 176
Db 121 IWENVVLIIVMACREFEMGRKKERYWPLYGEDPITPAFFKISCENEQARTDYFIRTL 180
Qy 177 KFNSETRIYGFHYKNWPDHVPSSIDPILELWVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 EFONESRLYGFHYVNWPDHVPSSDFSDILDMISLMRRYQEHEDVPICIHCSAGCGRTGA 240
Qy 237 ICAIVDYTWMLLKDGIIIPENSVFSLIREMTQRPVSLVOTQOYELVYNVNALEFKROMD 296
Db 241 ICAI-DYTWMLLKAGKIPDENVFNLIQEMTQRHSVOTKEQYELVHRAQAUFKEQLQ 299
Qy 297 VIRDKHSGTESQAKHCIPKNTLOADSYSPNLPKSTTTKAAKMMNQORTKMEIKESSFD 356
Db 300 -LYETHGAQKITDGNIEISTGNWSSIDSEKQDSPPP-----KPPRTRSLVEG----- 346
Qy 357 FRTSISAKKEELVLP 372
|||||:|

Db 347 -----DAKEEILQPP 356
RESULT 6
P70602
ID P70602 PRELIMINARY; PRT; 453 AA.
AC P70602;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE 20.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97067206; PubMed=8910608;
RA Aoki N., Yamaguchi-Aoki Y., Ullrich A.;
RT "The novel protein-tyrosine phosphatase PTP20 is a positive regulator
RT of PC12 cell neuronal differentiation.";
RL J. Biol. Chem. 271:29422-29426(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Aoki N., Yamaguchi-Aoki Y., Ullrich A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69673; AAC52896.1; -.
DR HSP; Q06124; 2SHP.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00700; PRTPHPHTASE.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 453 AA; 50106 MW; 6CCCC132206FB69AE CRC64;

Query Match 20.7%; Score 748; DB 11; Length 453;
Best Local Similarity 43.1%; Pred. No. 3.1e-45;
Matches 154; Conservative 58; Mismatches 117; Indels 28; Gaps 7;
Qy 3 QREILQFLDAQSKKITKEE-FANEFLLKLRQSTKYKADKTYPTVAENAKNIKKNRYK 61
Db 4 QSDLVRSFLEQEQEARDHRKGAILAREFSDIKARSVAMKTEGVCSTKAGSQGSKKNRYK 63
Qy 62 DILPYDSRVLSLITSDSDSSYINANFTKGVYGPAYATATOGPLSTTLTDFWRMWEYS 121
Db 64 DVVPYDETRVLSLLQEEGHGDIYINANFIRCTGSGYATATOGPLPHPTLLDFWRLVWFG 123
Qy 122 VLIIVMACMEYEMGKKKERYNAEPGEMQLEFGPFSVSCAEKR-KSDYIIRTLKVKNFS 180
Db 124 IKVILMACQETENGRRKRCERYNAEQERE-PLQAGPFCITLTRETALTSITURTLQVTFQK 182
Qy 181 ETRITVGFHYKNWPDHVPSSIDPILELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
Db 183 ESRPVHQLQYVMSWPDHGVPSSSDHLTMVEARCLQGLGPGPLCVHCSAGCGRTGVICA- 241
Qy 241 VDYTWMLLKDGIIIPENSVFSLIREMTQRPVSLVOTQOYELVYNVNALEFKROMD---- 296
Db 242 VDYVROLTLTOTIPNFSLFVWLEMRKQRPAAVQTEEQYRFLYHTVAQLFSRTLQNNSP 301
Qy 297 -----VIRDKHSGTESQAKHCIPKNTLOADSYSPNLPKSTTTKA 336
Db 302 LYQNLKENRAPICKDSSSLRTSSA---LPATSRPLGGVLRISVPGPPTLPMADTYA 355
RESULT 7
Q922E3
ID Q922E3 PRELIMINARY; PRT; 453 AA.
AC Q922E3;


```
Db 301 PHYONKENCAPICKEAFSLRTSSALP 327
:: :: || :: :: :: || :: ||
RESULT 9
Q99952 PRELIMINARY; PRT; 458 AA.
AC Q99952;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTEIN-TYROSINE-PHOSPHATASE (EC 3.1.3.48).
GN BDPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=87275816; PubMed=2956090;
RA Ralph S.J., Thomas M.L., Morton C.C., Trowbridge I.S.;
RT "Structural variants of human T200 glycoprotein (leukocyte-common
antigen).";
RL EMBO J. 6:1251-1257(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=85009812; PubMed=2971730;
RA Hall L.R., Streuli M., Schlossman S.F., Saito H.;
RT "Complete exon-intron organization of the human leukocyte common
antigen (CD45) gene.";
RL J. Immunol. 141:2781-2787(1988).
DR EMBL; M23492; AAD15273.2; -.
DR EMBL; M23496; AAD15273.2; JOINED.
DR EMBL; M23466; AAD15273.2; JOINED.
DR EMBL; M23467; AAD15273.2; JOINED.
DR EMBL; M23468; AAD15273.2; JOINED.
DR EMBL; M23469; AAD15273.2; JOINED.
DR EMBL; M23470; AAD15273.2; JOINED.
DR EMBL; M23471; AAD15273.2; JOINED.
DR EMBL; M23472; AAD15273.2; JOINED.
DR EMBL; M23473; AAD15273.2; JOINED.
DR EMBL; M23474; AAD15273.2; JOINED.
DR EMBL; M23475; AAD15273.2; JOINED.
DR EMBL; M23476; AAD15273.2; JOINED.
DR EMBL; M23477; AAD15273.2; JOINED.
DR EMBL; M23478; AAD15273.2; JOINED.
DR EMBL; M23479; AAD15273.2; JOINED.
DR EMBL; M23480; AAD15273.2; JOINED.
DR EMBL; M23481; AAD15273.2; JOINED.
DR EMBL; M23482; AAD15273.2; JOINED.
DR EMBL; M23483; AAD15273.2; JOINED.
DR EMBL; M23484; AAD15273.2; JOINED.
DR EMBL; M23485; AAD15273.2; JOINED.
DR EMBL; M23486; AAD15273.2; JOINED.
DR EMBL; M23487; AAD15273.2; JOINED.
DR EMBL; M23488; AAD15273.2; JOINED.
DR EMBL; M23489; AAD15273.2; JOINED.
DR EMBL; M23490; AAD15273.2; JOINED.
DR EMBL; M23491; AAD15273.2; JOINED.
DR HSSP; P18052; LYPO.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003595; PTPC_motif.
DR InterPro; IPR000387; TYR_phosphatase.
DR InterPro; IPR000242; TYR_prot_phptase.
DR Pfam; PF00041; fn3; 2.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00600; FN3; 2.
DR SMART; SM00194; PTPC; 2.
DR SMART; SM00404; PTPC_motif; 2.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 2.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.
KW Hydrolase.
FT NON_TER
FT SEQUENCE 1 1114 AA; 127625 MW; 055385D916234CC8 CRC64;
SQ SEQUENCE 1114 AA; 127625 MW; 055385D916234CC8 CRC64;

Query Match 19.6%; Score 708; DB 4; Length 458;
Best Local Similarity 46.7%; Pred. No. 2.2e-42;
Matches 136; Conservative 54; Mismatches 99; Indels 6; Gaps 4;

QY 25 ANEFLLKRSQYKADKTYPTTVAENAKNKKIKDILPYDYGRVELSLTSDSSY 84
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
DB 27 ACEFSDIQACSAWKADGVCSTVAGSRPENVRKNRYKDVLPYDQTRVILSLQEGHSDY 86
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
QY 85 INANIKGVYGPXKATQCGPLSTLLDFWRMIWEYSVLIIWMAWMEYEMGKKCERYWA 144
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
DB 87 INGNFIRGVGSLAYATQGPLPHTLLDFWRLWFEFGVKVILMACREIENGRKRCERYWA 146
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
QY 145 EPGEMOLEFGPFSVSCEAEK-RKSDYIITLKVKNSETRTIYQHYKNWPDHVPSSID 203
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
DB 147 QOQE-PLQTGLFCITLIEKWLNEIMLTLKVTQKESRVIQYQWSWPDGRGVPSSPD 205
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
QY 204 PILELIWVRVCYQEDSDVPICHCAGCGRTGCAIVDYTWMLLKDGIIPIENFSVSLI 263
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
DB 206 HMLAWVEEARLQGSPEPLCVCHCSAGCGRTGVLCT-VDYVRQLLTQMIPPDFSLFDV 264
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
QY 264 REMYRQPSLVTQOQYELVYNAVLELFRQMDVTRDKHSGTESAKHCIP 314
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||
DB 265 LKMRKORPAAVQTEQYRFLYHTVAQMF---CSTLQNASPHYQNIKENCAP 312
|| :: || :: || :: || :: || :: || :: || :: || :: || :: ||

RESULT 10
Q9H0Y6 PRELIMINARY; PRT; 1114 AA.
AC Q9H0Y6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T200 LEUKOCYTE COMMON ANTIGEN (FRAGMENT).
GN PTPRC.
```



```

DR MGD; MGI:104574; Ptpn1.
DR InterPro; IPR000387; Tyr_phosphatase.
DR Pfam; PF000242; Tyr_prot_phptase.
DR PRINTS; PR000700; PRTYPHPTASE.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
FT NON_TER 1
SQ SEQUENCE 361 AA; 41/726 MW; B51462E8D11023AA CRC64;

Query Match 14.2%; Score 513; DB 11; Length 361;
Best Local Similarity 39.08; Pred. No. 1.3e-28;
Matches 115; Conservative 50; Mismatches 110; Indels 20; Gaps 7;

Qy 15 QSKITKEEFANFLKLRQST-----KYKADK----TYPTTVAENAKNKNRYKDILP 65
Db 41 KSKLIRVENFEAYFKKQADSNCGFAEYEDKLIGISLPKYTAETAEARNRCKNRYNVLP 100

Qy 66 YDSRVLSLTSDSSYINANIKYGVGPAYATQGPLSTLLDFWRMIWEYSVLII 125
Db 101 YDISRVLSVQTHSTD-DYINANTYMPGVHKKDFIATQGPLNTLKDFWRMVWEKNYAI 159

Qy 126 VMACMEYEMGKKKERYWAEPGEMQLEFPFVSCEAEKRSKYIIRTLKVK--FNSETR 183
Db 160 VMLTKVCEQGRKCEEYW--PSKAQDYGDTIVANTSEVLPEWTIRDFVYKNQNSSEH 217

Qy 184 TYQFHYKNWPDHDPSSIDPILSLINDVRCY--QEDDSVPICHCAGCGRTGVCIAIV 241
Db 218 PLRQFHFTSPDHGVPDITDLLINFRYLVRDYMKQIPPEPILVHCSAGVGRGTGTFIAD 277

Qy 242 DYTWMLLKDGILPENFSVFLSIREMRTORPSLVQTOEQYELVYNVLELFRQMD 296
Db 278 RLIIQIENENTV---DVGIVYDLRMLRPLMVQTEDQYVFLNQCVLDDIIRAQKD 328

RESULT 13
Q9PUM2 PRELIMINARY; PRT; 1195 AA.
AC Q9PUM2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RECEPTOR PROTEIN TYROSINE PHOSPHATASE RHO PRECURSOR (FRAGMENT).
GN PTPRT.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20193505; PubMed=10727868;
RA Johnson K.G., Holt C.E.;
RT "Expression of CRYP-alpha, LAR, PTP-delta, and PTP-rho in the
RL developing xenopus visual system.";
RL Mech. Dev. 92:291-294(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Besco J.A., Johnson K.G., Frostholt A., Popesco M., Burghes A.H.M.,
RA Holt C.E., Rotter A.;
RT "Xenopus laevis receptor protein tyrosine phosphatase rho.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AF173857; AAD50295.1; -.
DR HSSP; P28827; 1RPM.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000998; MAM.
DR InterPro; IPR000387; Tyr_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.

```

```

DR Pfam; PF00041; fn3; 3.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF00102; Y_phosphatase; 1.
DR PRINTS; PR00020; MAMDOMAIN.
DR PRINTS; PR00700; PRTYPHPTASE.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00137; MAM; 1.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00740; MAM_1; UNKNOWN_1.
DR PROSITE; PS00060; MAM_2; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
KW Glycoprotein; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 >1195 RECEPTOR PROTEIN TYROSINE PHOSPHATASE
FT NON_TER 1195 1195 RHO.
FT SQ SEQUENCE 1195 AA; 132818 MW; 71F3C6D0AAD34F07 CRC64;

Query Match 14.1%; Score 511; DB 13; Length 1195;
Best Local Similarity 41.5%; Pred. No. 8.7e-28;
Matches 107; Conservative 47; Mismatches 92; Indels 12; Gaps 4;

Qy 37 KYKA---DKTYPTTVAENAKNKNRYKDILPDYDSRVLSLTSDSSYINANIKGV 93
Db 908 EYEALPEQPTASDTAKEDENRNKNRYGNIISYDHSRVRLQIEEDPHSDYINANYVDGY 967

Qy 94 YGPKAYATQGPLSTLLDFWRMIWEYSVLIIWACMEYEMGKKKERYWAEPGEMQLEF 153
Db 968 HRPRYIATQGPMDQETVKDFWRMIWQENSISVVMVTNLVEVGRVKCVWPDDTEV---Y 1024

Qy 154 GPFVSCEAEKRSKYIIRTLKVK--KENSEPTTYQFHYKNWPDHDPSSIDPILSLIND 211
Db 1025 GDIKVSATETPLAEYIIRTTVQKGSHEIREIRQHFHTSWPDNGVPCYATGLGFIRQ 1084

Qy 212 VRCYQEDDSVPICHCAGCGRTGVCIAIVDTYTWMLLKDGILPENFSVFLSIREMRTQRP 271
Db 1085 VKFLNPPDAGPITVHCSAGAGRTGCFIAIDIMLDMAEKGVV---DIFNCVRELRAQRV 1140

Qy 272 SLVQTOEQYELVYNVLE 289
Db 1141 NMVQTEEQYVFEHDAILE 1158

RESULT 14
Q62884 PRELIMINARY; PRT; 1216 AA.
AC Q62884;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DENSITY-ENHANCED PHOSPHATASE-1 PRECURSOR (EC 3.1.3.48) (DEP-1)
DE (VASCULAR PROTEIN TYROSINE PHOSPHATASE 1).
GN DEPI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR KYOTO; TISSUE=AORTA;
RX MEDLINE-96375109; PubMed=8781490;
RA Borges L.G., Selfert R.A., Grant F.J., Hart C.E., Distèche C.M.,
RA Edeihoff S., Solca F.F., Lieberman M.A., Lindner V., Fischer E.H.,
RA Lok S., Bowen-Pope D.F.;
RT "Cloning and characterization of rat density-enhanced phosphatase-1, a
RL protein tyrosine phosphatase expressed by vascular cells.";
RL Circ. Res. 79:570-580(1996).
CC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND/OR SIGNALING EVENTS
CC INVOLVING CELL-CELL AND CELL-MATRIX CONTACT.
CC -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN

```


CC		TYROSINE + PHOSPHATE.
CC	-1-	SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC	-1-	TISSUE SPECIFICITY: HIGH EXPRESSION IN THE CEREBELLUM, BRAIN CORTEX, AND KIDNEY CORTEX, AND SOMEWHAT LESS ABUNDANT IN SPLEEN AND LUNG.
CC	-1-	SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC	-1-	SIMILARITY: CONTAINS 1 PROTEIN-TYROSINE PHOSPHATASE DOMAIN.
DR	EMBL:	U40790; AAB53195.1; -.
DR	HSSP:	P18052; LYFO.
DR	InterPro:	IPR003961; FN_III.
DR	InterPro:	IPR000387; TYR_phosphatase.
DR	InterPro:	IPR000242; Tyr_prot_phphatase.
DR	Pfam:	PF00041; fn3; 7.
DR	Pfam:	PF00102; Y_phosphatase; 1.
DR	PRINTS:	PR00700; PRTPHPHTASE.
DR	SMART:	SMW0060; FN3; 5.
DR	SMART:	SMW0194; PTPC; 1.
DR	PROSITE:	PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR	PROSITE:	PS0056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE:	PS00555; TYR_PHOSPHATASE_Ptp; 1.
KW	Signal:	Glycoprotein; Transmembrane; Repeat; Hydrolase; Receptor.
FT	SIGNAL	1-28 POTENTIAL.
FT	CHAIN	29-1216 DENSITY-ENHANCED PHOSPHATASE-1.
FT	DOMAIN	21-28 POLY-LEU.
FT	DOMAIN	29-854 EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	695-698 POLY-PRO.
FT	TRANSMEM	855-875 POTENTIAL.
FT	DOMAIN	876-1216 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	39-102 FIBRONECTIN TYPE-III.
FT	DOMAIN	122-225 FIBRONECTIN TYPE-III.
FT	DOMAIN	245-314 FIBRONECTIN TYPE-III.
FT	DOMAIN	333-359 FIBRONECTIN TYPE-III.
FT	DOMAIN	419-486 FIBRONECTIN TYPE-III.
FT	DOMAIN	505-575 FIBRONECTIN TYPE-III.
FT	DOMAIN	597-672 FIBRONECTIN TYPE-III.
FT	DOMAIN	701-787 FIBRONECTIN TYPE-III.
FT	DOMAIN	940-1216 PROTEIN-TYROSINE PHOSPHATASE.
FT	ACT_SITE	1118-1118 BY SIMILARITY.
FT	CARBOHYD	62-62 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	83-83 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	88-88 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	108-108 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	118-118 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	147-147 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	162-162 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	184-184 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	256-256 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	261-261 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	271-271 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	291-291 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	295-295 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	311-311 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	344-344 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	357-357 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	376-376 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	405-405 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	459-459 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	465-465 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	483-483 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	496-496 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	508-508 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	517-517 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	549-549 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	640-640 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	646-646 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	651-651 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	663-663 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	669-669 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	703-703 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	789-789 N-LINKED (GLCNAC..) (POTENTIAL).
FT	CARBOHYD	816-816 N-LINKED (GLCNAC..) (POTENTIAL).
SQ	SEQUENCE	1216 AA; 134275 MW; 4D6193B108E53591 CRC64;

QY 66 YDYSRVELSLITSDSSYINANFIKGYGPKAYIATQGPLSTTLDDFWRMIWEYSVLII 125
Db 1077 YDISRVKLSVQTHSTD-DYINANYMPGVHSHKKDFIATQGPLNTLKDFRWWKKNVYAI 1135
QY 126 VMACMEYEMGKKCBRYNAEPGEMOLEFGPFSVSCAEKRSYIIRTLKVK--FNSETR 183
Db 1136 IMLTKCVEQGRTKCEYW--PSKQADYGDITVAMTSEIVLPFWTIRDTYKNIQTSSEH 1193
QY 184 TIYQFHYKNWPDHVPSSIDPILWIDVRCY--QEDDSVPICIHCSAGCGRTGVICAIV 241
Db 1194 PLRQFHTSWPDHGVDPDTDLLINFRYLVRDYMKOSPESPILVHCSAGVGRGTGTFIAD 1253
QY 242 DYTWMLLKDGIIPENSVFSFIREMRTORPSLVQTEQYELVYNNAVLELFKQMD 296
Db 1254 RLIIYQIENENTV----DVYGIYVDLRMHRPLMVQTEQYVFLNQCVDIVRSQKD 1304

Search completed: May 26, 2002, 12:26:01
Job time: 546 sec

OM of: US-09-600-358a-4 to: EST:* out_format : pfs

Date: May 26, 2002 1:04 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet p2n.model -DEV=xlp
-O=/cgn2.1/USPQC_spool/US09600358/runat_24052002_101553_26989/app_query.fasta_1.761
-DB=EST -QFMT=fastap -SUFFIX=first -GAPOF=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOF=4.500
-GAPEXT=0.050 -XGAPOF=10.000 -XGAPEXT=0.500 -FGAPOF=6.000
-FGAPEXT=7.000 -YGAPOF=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-ALIGN=7 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-LIST=45 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09600358.@CGN1_1_7299
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-600-358a-4

Query length: 692

Database: EST*

Database sequences: 13736207

Database length: -1841457050

Search time (sec): 2277.120000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_hlc:AF077031	+	3512.50	5975.82	0.0	2903	! AF077031 Homo sapiens protein
gb_est2:BM458477	+	1578.00	2675.89	8.4e-140	1116	! BM458477 AGENCOURT_6413950 NIH
gb_est2:BG541535	+	1112.00	1881.73	1.4e-95	832	! BG541535 602570982F1 NIH_MGC_77
gb_est1:AU132052	+	929.50	1569.87	3.4e-78	796	! AU132052 AU132052 NT2RP3 Homo s
gb_est2:BM469079	+	887.00	1494.34	5.5e-74	1003	! BM469079 AGENCOURT_6445886 NIH
gb_est1:AN967600	+	868.50	1469.21	1.4e-72	582	! AN967600 EST379675 MAGE resequ
gb_est1:AW504119	+	848.50	1437.42	8.1e-71	475	! AW504119 UI-HF-BN0-all-h-10-0-0
gb_est1:AU132040	+	835.00	1407.88	3.6e-69	811	! AU132040 AU132040 NT2RP3 Homo s
gb_est1:AW373463	+	824.00	1394.37	2.0e-68	521	! AW373463 RC0-T0522-071299-011
gb_est1:AW363814	+	788.00	1331.29	6.6e-65	588	! AW363814 RC0-T0358-071299-011
gb_est1:AW915666	+	760.00	1284.08	2.8e-62	554	! AW915666 EST346970 Normalized x
gb_est2:BG498752	+	738.00	1242.86	5.6e-60	745	! BG498752 602544381F1 NIH_MGC_60
gb_est2:BL152499	+	718.50	1208.11	4.8e-58	835	! BL152499 602915640F1 NCI_CGAP_1
gb_est1:AW822729	+	694.50	1170.43	6.0e-56	629	! AW822729 uq17g09.y1 Ren Scubbs
gb_est2:BM194577	+	661.50	1116.58	6.0e-53	505	! BM194577 TCAAP3E10089 Pediatr ic
gb_est2:BI411856	+	646.00	1081.20	5.6e-51	1054	! BI411856 602966483F1 NCI_CGAP_1
gb_est1:AJ394015	+	621.50	1041.72	8.9e-49	859	! AJ394015 AJ394015 dkf426 Gallu
gb_est1:AL597518	+	611.00	1031.58	3.3e-48	448	! AL597518 DKZP313F0814_r1 313 0
gb_est2:BF301903	+	571.00	954.30	6.6e-44	932	! BF301903 602032931F1 NCI_CGAP_1
gb_est1:AA851710	+	569.00	962.48	2.3e-44	355	! AA851710 EST194478 Normalized x
gb_est2:BF105998	+	561.00	938.94	4.7e-43	805	! BF105998 601823023F1 NIH_MGC_77
gb_est1:AI530640	+	556.00	930.78	1.3e-42	836	! AI530640 mp70q01.y1 Soares thym
gb_est2:BE765069	+	550.00	925.84	2.5e-42	500	! BE765069 CMI-NT0087-280600-291-
gb_est2:BG614041	+	545.00	911.42	1.6e-41	814	! BG614041 603641583F1 NIH_MGC_61
gb_est2:BG505436	+	539.00	904.42	3.9e-41	620	! BG505436 602549331F1 NIH_MGC_61
gb_est2:BI415856	+	533.50	891.72	2.0e-40	815	! BI415856 602988052F1 NCI_CGAP_1
gb_est1:AA998917	+	530.00	896.60	1.1e-40	330	! AA998917 UI-R-C-02-0-01.s
gb_est1:AL522809	+	508.50	851.10	3.7e-38	680	! AL522809 AL522809 LTI_NFL004.NE
gb_est1:AL600035	+	502.00	845.19	7.9e-38	441	! AL600035 DKZP31302031_r1 313 0
gb_est2:BI691928	+	497.00	830.56	5.1e-37	730	! BI691928 603311861F1 NCI_CGAP_M
gb_est1:AU121737	+	485.00	810.16	7.0e-36	722	! AU121737 NT2RP2 Homo s
gb_est2:BF099411	+	485.00	807.33	1.0e-35	913	! BF099411 601751512F1 NCI_CGAP_M
gb_est1:AI323259	+	481.50	808.35	8.9e-36	510	! AI323259 mp91b04.y1 Soares thym
gb_est2:BI415322	+	479.50	795.79	4.4e-35	1090	! BI415322 602987967F1 NCI_CGAP_1
gb_est1:AA241226	+	477.00	798.47	3.1e-35	611	! AA241226 mu95d08.r1 Soares mous
gb_est2:BF579759	+	476.00	791.09	8.1e-35	979	! BF579759 602095070F1 NCI_CGAP_C
gb_est2:BG694807	+	472.50	788.88	1.1e-34	715	! BG694807 602829360F1 NCI_CGAP_C
gb_est2:BE765067	+	455.00	765.29	2.2e-33	421	! BE765067 CMI-NT0087-280600-291-
gb_est1:AI981941	+	454.00	759.83	4.5e-33	575	! AI981941 pat.pk0069.d6.f chicke
gb_est2:BI686558	+	451.00	749.41	1.7e-32	892	! BI686558 603313654F1 NCI_CGAP_M

gb_est2:BF144494 + 448.50 747.71 2.1e-32 720 ! BF144494 601790135F1 NCI_CGA
gb_est2:BG234079 + 447.50 754.14 9.3e-33 366 ! BG234079 daa36e12.y1 NICHDX
gb_est1:AW766206 + 447.00 746.13 2.6e-32 663 ! AW766206 da83f09.y1 Harland
gb_est1:AL555127 + 446.00 743.40 3.7e-32 722 ! AL555127 AL555127 LTI_NFL006
gb_est1:AW319424 + 442.00 737.20 8.1e-32 684 ! AW319424 un16c07.y1 Sugano m

seq_name: gb_hlc:AF077031

seq_documentation_block:

LOCUS AF077031 2903 bp mRNA linear HTC 22-MAY-2001
DEFINITION Homo sapiens protein tyrosine phosphatase homolog mRNA, complete cds.

ACCESSION AF077031 GI:4689109

VERSION AF077031.1

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2903)

AUTHORS Liu,T., Zhang,J., Fu,G., Zhang,Q., Ye,M., Zhou,J., Wu,J., Shen,Y.,
Yu,M., Chen,S., Mao,M. and Chen,Z.

TITLE Human protein tyrosine phosphatase (70zpep) homolog

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2903)

AUTHORS Liu,T.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-1998) Shanghai Institute of Hematology, Shanghai
Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II,
Shanghai 200025, P.R. China

FEATURES

Location/Qualifiers

1..2903

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="NB4"

73..2496

/codon_start=1

/product="protein tyrosine phosphatase homolog"

/protein_id="AAD27764.1"

/db_xref="GI:4689110"

/translation="MDREILQFLDEAQSQKITKEEFANEFKLKLRQSTKYKADKTY
PTAAATKPKNKKRYKDIPIYDSRVSLTSDESSYINANIKGYIGPKAYIAT
OQPLSTLLDFWRMTKSVLIIGMACMEYKKCYWAEYVPELEFGPFSVC
EAEKRSYIITLAVKNSRTIYQFHYKNWPDHVPSSIDPILWVRCVQED
DSVPCIHCSACGRFGICADITWMLLKDGLIENSVSLIREMTQRSLVQTQ
EYELVYNAVLELFRQMDVIRKDHSGTESQAHCIPERNHTFLQADSPNLPKSTTK
AAKMNQORTKMEIKESSDFRTSEISAKELVLHPAKSTSFDFLELNYSFKNAD
TMMKQTKAFPIVGPGLQKHSGLDGLSGFEGCSNKPVNAAGFYNSKVPITRTKST
PPELQOQRETKVDSENFYSLESDPHDSCFVEMQAQKVMHVSAAELNYSLPYDSKHQ
IRNASNVKHDSSALGVSYIPLVENPYESSWPPSGTSKKNSLDLPEKODGVFPVPSL
LPTSTSLFYSYNSHDSLSLSNPTNISLLNGESAVLATAPRIDDEIPPLPFWTPES
FIVVEAEGRSPNPKSVASAVKVIKIGTSLWGGTSEPKKFDSDVILRPSKVLKLRSP
KSELHQDRSSPPPLPRTLEFSLFLEADCEMQAQSIETYSYPTDMENSTSSKQTLK
TSGLKFTSRSKSLILRNKKSCICNCPNPAESVQSNSSSFLNFGNFRPSKPGP
RNPPTWNI"

BASE COUNT 955 a 580 c 528 g 840 t

ORIGIN

alignment_scores:

Quality: 3512.50

Ratio: 5.173

Percent Similarity: 99.124

Length: 685

Gaps: 1

Percent Identity: 98.832

alignment_block:

US-09-600-358A-4 x AF077031 ..

Align seg 1/1 to: AF077031 from: 1 to: 2903

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
73 ATGACCAAGAGAAATCTCGAGAGTCTCTGATGAGGCCCAAGCAA 122

```
17 slyslleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArgG 34
|||||
123 GAAATTTACTAAAGAGAGCTTTGCCAATGAATTTCTGAAGCTGAAAGGC 172
34 lnSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
173 AATCTACCAAGTACAAGGAGAGACAAACCTATCCTACAACTGTGGCTGAG 222
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProThrAs 67
::: |||||
223 AGCCCAAGAANTATCAAGAAAAACAGATATAGGATATTTGGCCCTATGA 272
67 pTyrSerArgValGluSerLeuIleThrSerAspGluAspSerSerT 84
|||||
273 TTATAGCCGGGTAGAACTATCCCTGATACCTCTGATGAGGATCCAGCT 322
84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
323 ACATCAATGCCAACTTCATTAGGGAGTTTATGGACCCAAGGCTTATATT 372
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
|||||
373 GCCACCCAGGGTCTTTATCTACAACCCCTCTGGACTTCTGGAGGATGAT 422
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
423 TTGGGAATATAGTCTCTTATCATTTGGTATGGCATGCATGGAGTATGAAA 472
134 etGlyLysLysCysGluArgTyrTyrTrpAlaGluProGlyGluMetGln 150
|||||
473 TGGGAAGAAAAAGTGTGAGCGCTACTGGGCTGAGCCAGTAGAGATGCCAG 522
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
523 CTGGAAATTTGGCCCTTCTCTGTATCCTGTGAAGCTGAAAAAGGAAATC 572
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
573 TGATTTATATAATCAGGACTCTAAAAGTTAAAGTTCAATAGTGAACTCGAA 622
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisaspValProSer 200
|||||
623 CTATCTACCAAGTTTCATTACAAGAAATGGCCAGACCATGATGTACCTTCA 672
201 SerIleAspProIleLeuGluLeuIleTyrAspValArgCystyrGlnGl 217
|||||
673 TCTATAGACCTTATCTTGAGCTCATCTGGGATGTACGTTGTTACCAAGA 722
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
723 GGATGACAGTGTCCCATATGCATTCACATGCAGTGTGGCTGTGGAAGGA 772
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
773 CTGGTGTATTGTGTATT...GATTATACATGGATGTGTCTAAAGAT 819
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetar 267
|||||
820 GGGATAAATTCCTGAGAACTTCAGTGTTTTTCAGTTTGTATCCGGGAAATGCG 869
267 gThrGlnArgProSerLeuValGlnThrGlnGluInTyrGluLeuValIt 284
|||||
870 GACACAGAGGCCCTTCATTAGTTCAACAGCGCAGGAACAATATGAACCTGGTCT 919
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
920 ACAATGCTGTATTAGAACTATTTAAGACAGACAGATGGATGTTATCAGAGAT 969
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||
970 AAMCATTTCTGGAACAGAGATCAAGCAAGCAATGTATTCTCTCAGAAAAA 1019
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
|||||
```

```
1020 TCACACTCTCCAAGCAGACTCTTATTCTCTTAATTTACCAAAAAGTACCA 1069
334 hrLysAlaAlaLysMetMetAsnGlnGlnArgThrLysMetGluIleLys 350
|||||
1070 CAAAAGCAGCAAAAATGATGAAACCAACAAGGACAAAAATGAAATCAAA 1119
351 GluSerSerSerPheAspPheArgThrSerGluIleSerAlaLysGluGl 367
|||||
1120 GAATCTTCTCTCTTTGACTTTAGGACTTCTGAATAAGTGCAAAAGAGA 1169
367 uLeuValLeuHisProAlaLysSerThrSerPheAspPheLeuGluL 384
|||||
1170 GCTAGTTTTGCAACCCTGCTAAATCAAGCACTTCTTTTGACTTTCTGGAGC 1219
384 euAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnThr 400
|||||
1220 TAAATTACAGTTTTTGACAAAAATGCTGACACAACCATGAAATGGCAGACA 1269
401 LysAlaPheProIleValGlyGluProLeuGlnLysHisGlnSerLeuAs 417
|||||
1270 AAGGCATTTCCAAATAGTTGGGAGCCTCTTCAGAGCATCAAAGTTTGG 1319
417 pLeuGlySerLeuLeuPheGluGlyCysSerAsnSerLysProValAsnA 434
|||||
1320 TTTGGGCTCTCTCTTTGTTGAGGGATGTTCTAATTTCTAAACCTGTAAATG 1369
434 laAlaGlyArgTyrPheAsnSerLysValProIleThrArgThrLysSer 450
|||||
1370 CAGCAGGAAGATATTTTAAATCAAAAGTGCCAAATAACACGACCAATCA 1419
451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
|||||
1420 ACTCTTTTGAATTTGATACAGCAGAGAGAAACCAAGAGGTGGACAGCAA 1469
467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
|||||
1470 GGAATAACTTTTCTTATTGGAATCTCAACACATGATCTCTGTTTGTAG 1519
484 luMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
|||||
1520 AGATGAGGCTCAAAAAGTAATGCATGTTCTTCACGACAACCTGAATAT 1569
501 SerLeuProTyrAspSerLysHisGlnIleArgAsnAlaSerAsnVally 517
|||||
1570 TCACTGCCATATGACTCTAAACACCAAAATACGTAATGCCTCTAATCTAAA 1619
517 sHisHisAspSerSerAlaLeuGlyValTyrSerTyrIleProLeuValG 534
|||||
1620 GCACCATGACTCTAGTCTCTTTGGTGTATATTCTTACATACCTTTTAGTG 1669
534 luAsnProTyrPheSerSerTrpProSerGlyThrSerSerLysMet 550
|||||
1670 AAAATCCTTATTTTTCATGCGCTCCCAAGTGGTACCAGTTCTTAAGATG 1719
551 SerLeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLe 567
|||||
1720 TCTCTTGATTTACCTGAGAAGCAAGATGGAACGTGTTTTTCTCTCTCTCT 1769
567 uLeuProThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerS 584
|||||
1770 GTTGCCAAACATCCTCTACATCCCTCTCTTCTTATTACAAATTCACATGAT 1819
584 erLeuSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGlu 600
|||||
1820 CTTTATCACTGAATTTCTCAACCAATATTTCTCTACTATTGAACAGGAG 1869
601 SerAlaValLeuAlaThrAlaProArgIleAspAspGluIleProPropr 617
|||||
1870 TCAGCTGTACTAGCAACTGCTCCAAGATAGATGATGAAATCCCCCTCC 1919
617 oLeuProValArgThrProGluSerPheIleValValGluAlaLysGlyC 634
|||||
```

```

1920 ACTTCCTGTATGGACACCTGAATCATTTATGTTGTTGAGGAAGCTGGAG 1969
634 luPheSerProAsnValProLysSerLeuSerAlaValLysValLys 650
|||||
1970 AATTCTCACCACAAATGTTCCCAATTCCTTATCTCAGCTGTGAGGTAAHA 2019
|||||
651 lIeGlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAs 667
|||||
2020 ATTGGAACATCACTGGAATGGGTGGAACATCTGAACCAAGAAATTTGA 2069
|||||
667 pAspSerValIleLeuArgProSerLysSerValLysLeuArgSerProL 684
|||||
2070 TGACTCTGTGATACTTAGACCAAGCAAGAGTGAATAACTCCGAAGTCTCA 2119
|||||
684 ysSer 685
|||||
2120 AATCA 2124

```

seq_name: gb_est2:BM458477

```

seq_documentation_block:
LOCUS BM458477 1116 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6413950 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497108
5', mRNA sequence.

```

```

ACCESSION BM458477
VERSION BM458477.1 GI:18507517
KEYWORDS EST.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1116)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

```

```

Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: LLAM12126 row: 1 column: 05
High quality sequence stop: 717.
Location/Qualifiers

```

FEATURES

```

source
1..1116
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5497108"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

```

BASE COUNT 350 a 232 c 232 g 302 t
ORIGIN

```

alignment_scores:

```

Quality: 1578.00 Length: 404
Ratio: 4.767 Gaps: 3
Percent Similarity: 81.931 Percent Identity: 78.960

```

alignment_block:

US-09-600-358A-4 x BM458477 ..

Align seg 1/1 to: BM458477 from: 1 to: 1116

```

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
75 ATGGACCAAAAGAGAAATTCGACAGAGTTCCTGATGAGGCCCAAGCAA 124
|||||
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
|||||
125 GAAATTAATAAGAGGAGTTTGCCAAATGAATTTCTGAAGCTGAAAGGC 174
|||||
34 lnsrThrLysThrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
|||||
175 AATCTACCAAGTACAGGACAGACAAACCTATCTACAACTGTGGCTGAG 224
|||||
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
::: |||||
225 AAGCCCAAGAATATCAAGAAAAACAGATATAAGGATATTTTGCCTATGA 274
|||||
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84
|||||
275 TTATAGCCGGGTAGAACTATCCCTGATAACCTCTGATGAGGATTCACAG 324
|||||
84 yrlleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
|||||
325 ACATCAATGCCAATTCATTAAAGGAGTTTATGACCAAGGCTTATATT 374
|||||
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTyrArgMetIl 117
|||||
375 GCCACCCAGGGTCTCTTATCTACAACCTCTGGACTTCTGGAGGATGAT 424
|||||
117 eTrpGluTyrSerValLeuIleIleValMetAlaCysMetGluTyrGluM 134
|||||
425 TTGGGAATATAGTCTCCTTATCATTTGTTATGGCATGATGAGTATGAAA 474
|||||
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
|||||
475 TGGAAAGAAAAGTGGAGCGCTACTGGGCTGAGCCAGAGAGATGCGAG 524
|||||
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysSe 167
|||||
525 CTGGAATTTGGCCCTTCTCTGTATCTCTGAAGCTGAAAAAGGAATC 574
|||||
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThrArgT 184
|||||
575 TGATTATATATATCAGGACTCTAAAAGTTAAGTTCAATAGTGAACACTCAA 624
|||||
184 hrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValProSer 200
|||||
625 CTATCTACCAAGTTTCATTACAAGAAATGGCCAGACCATGATGATACCTTCA 674
|||||
201 SerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrGlnGl 217
|||||
675 TCTATAGACCTATTCCTGAGCTCATCTGGGATGATGCTTGTACCAAGA 724
|||||
217 uAspAspSerValProIleCysIleHisCysSerAlaGlyCysGlyArgT 234
|||||
725 GGATGACAGTGTCTCCATATGCAATTCAGTCTGCTGCTGCTGGAAGGA 774
|||||
234 hrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLysAsp 250
|||||
775 CTGGTGTATTGTCGCTATT...GATTATACATGATGATGTTGCTAAAGAG 821
|||||
251 GlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluMetAr 267
|||||
822 GGG..... 824
|||||
267 gThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeuValT 284
|||||
824 ..... 824
|||||
284 yrAsnAlaValLeuGluLeuPheLysArgGlnMetAspValIleArgAsp 300
|||||
824 ..... 824
|||||
301 LysHisSerGlyThrGluSerGlnAlaLysHisCysIleProGluLysAs 317
|||||

```

825AGTCAAGCCAAAGCATTTGTTATTCCTGAGAAAA 856
317 nHisThrLeuGlnAlaAspSerTyrSerProAsnLeuProLysSerThrT 334
857 TAC.ACTCCTCAGCAGACTCTTATCTCTATTTACCAAAAGTACCC 905
334 hrLysAlaAlaLysMetMetAsnGlnArgThrLysMetGluLys 350
906 CAAGCCGCAAAATTTGATGAACCAACCAAGGACAAATGGAAATCAA 955
351 GluSerSerSer.PheAspPheArgThrSerGluLysSerAlaLysGluG 367
956 GAATCTCTCTCCCTTTGACTTTTAGACTCTCTGAATAGG.GCACAAGAAG 1004
367 luleuValLeuHsProAlaLysSerSerThrSerPheAspPheLeuGlu 383
1005 AGTAGTTTTCACCTGCTTAATCAGGACTTCTTTTGTGCTTCTGGA 1054
384 LeuAsnTyrSerPheAspLysAsnAlaAspThrThrMetLysTrpGlnTh 400
1055 GCAAAATTACGTTTGTGAATAATGCTGAACCAACTTGAATGGCCGAA 1104
400 rLysAlaPhe 403
1105 CAAGGCATTT 1114

seq_name: gb_est2:BG541535

seq_documentation_block:
LOCUS BG541535 832 bp mRNA linear EST 03-APR-2001
DEFINITION 602570982P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4695140 5',
mRNA sequence.

ACCESSION BG541535
VERSION BG541535.1 GI:13533768
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 832)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov.
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI520 row: d column: 21
High quality sequence stop: 785.
Location/Qualifiers
1. 832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4695140"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

FEATURES
Source

BASE COUNT 245 a 198 c 136 g 252 t 1 others

ORIGIN
alignment_scores:
Quality: 1112.00 Length: 223
Ratio: 5.032 Gaps: 1
Percent Similarity: 99.103 Percent Identity: 98.206
alignment_block:
US-09-600-358A-4 x BG541535 ..
Align seg 1/1 to: BG541535 from: 1 to: 832
464 ValAspSerLysGluAsnPheSerTyrLeuGluSerGlnProHisAspSe 480
1 GTGGACAGCAAGAAACTTTTCTTATTTGGAACTCTCAACACACATGATTC 50
480 rCysPheValGluMetGlnAlaGlnLysValMetHisValSerSerAlag 497
51 TTGTTTTGTAGAGATGCAGGCTCAAAAGTAATGCATGTTTCTTCAGCAG 100
497 luleuAsnTyrSerLeuProTyrAspSerLysHisGlnIleArgAsnAla 513
101 AACTGAATTTACTACTGCCATATGACTCTAAACACCAAAATACGTAATGCC 150
514 SerAsnValLysHisHisAspSerSerAlaLeuGlyValTyrSerTyrI1 530
151 TCTAATGTAAAGCACCACCTAGCTCTAGTGTCTTGGTGATATATCTTACAT 200
530 eProLeuValGluAsnProTyrPheSerSerTrpProSerGlyThrs 547
201 ACCTTTAGTGAATAATCCCTATTTTTCATCATGCGCTCCCAAGTGGTACCA 250
547 erSerLysMetSerLeuAspLeuProGluLysGlnAspGlyThrValPhe 563
251 GTTCTAAGATGCTCTTGATTTACCTTGAGAGCAAGATGGAATGTTTT 300
564 ProSerSerLeuLeuProThrSerSerThrSerLeuPheSerTyrTrAs 580
301 CCTTCTCTCTGTGGCAACATCCTCTACATCCCTCTCTCTATTATACAA 350
580 nSerHisSerSerLeuSerLeuAsnSerProThrAsnIleSerSerLeuL 597
351 TTCACATGATCTTTTATCATCTGAATTTCTCAACCAATATTTCTCTACTAT 400
597 euAsnGlnGluSerAlaValLeuAlaThrAlaProArgIleAspaspGlu 613
401 TGAACCCAGGAGTCAGCTGTACTAGCAACTGCTCCCAAGGATAGATGA 450
614 IleProProLeuProValArgThrProGluSerPheIleValValG1 630
451 ATCCCCCTCCACTTCTCTGACGGACACCTGAATCATTTATNGTGGTGA 500
630 uGluAlaGlyGluPheSerProAsnValProLysSerLeuSerSerAlav 647
501 GGAAGCTGGAGAAATCTCACAAATGTTCCCAAAATCCTTTATCTCAGCTG 550
647 alLysValLysIleGlyThrSerLeuGluTrpGlyGlyThrSerGluPro 663
551 TGAAGTAAATAATTTGGAACATCACTCTGGAATGGGGTGGAAACATCTGA 600
664 LysLysPheAspAspSerValIleLeuArgProSerLys.SerValLysL 680
601 AGAATAATTGATGACTCTGTGATACTTTAGACCAAGCAACGAGTGTAAAC 650
680 euArgSerProLysSer 685
651 TCCGAAGTCTCTAAATCA 667
seq_name: gb_est1:AU132052
seq_documentation_block:
LOCUS AU132052

796 bp mRNA linear EST 24-OCT-2000

```

DEFINITION AUI32052 NT2RP3 Homo sapiens cDNA clone NT2RP3003698 5', mRNA
sequence.
ACCESSION AUI32052
VERSION AUI32052.1 GI:10992406
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 796)
AUTHORS Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
Isogai.T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3003698"
/clone_lib="NT2RP3"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 281 a 132 c 162 g 217 t 4 others
ORIGIN

alignment_scores:
Quality: 929.50 Length: 254
Ratio: 4.187 Gaps: 1
Percent Similarity: 87.402 Percent Identity: 66.535

alignment_block:
US-09-600-358a-4 x AUI32052 ..
Align seg 1/1 to: AUI32052 from: 1 to: 796

33 ArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThrValAl 49
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 AGATTGCTACCAATAATAGACAAGAAAGATATATCCACAGCCACTGG 50
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
49 aGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProT 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 AGAAAGAGACAAATGTTAAAGACACAGATACAGACATAGTGGCAT 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 yRAspTyrSerArgValGluLeuSerLeuThrSerAspGluAspSer 82
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 TTGATCACAGCCGAGTTAAATTTGACATTAAGACTCTTCACAGATTCA 150
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 SerTyrIleAsnAlaAsnPhelIleLysGlyValTyrGlyProLysAlaTy 99
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
151 GACTATATCATGCAATTTTATAAGGGCTCTATGGGCCAAAGGATA 200
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 rIleAlaThrGlnGlyProLeuSerThrThrLeuAspPheTrpArgM 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
201 TGTAGCAACTCAGGACCTTAGCAATACAGTAATAGATTTTGGAGGA 250
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 etIleTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyr 132
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
251 TGATATGGAGTATAATGTTGTGATCATTTGTAATGGCTGCCGGAATTT 300

```

```

133 GluMetGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluLe 149
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GAGATGGGAAGAAAATATGTGAGCGCTATGGCTTTGTATGGAGAAGA 350
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 tGlnLeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgL 166
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 CCCATACACGTTTGCACCATTTAAATTTCTTGTGAGGATGAACAAGCAA 400
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
166 ySerAspTyrIleIleArgThrLeuLysValLysPheAsnSerGluThr 182
::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
401 GAACAGACTACTTCTCAGGACACTTACTTGAATTTCAAAATGAATCT 450
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 ArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisAspValPr 199
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
451 CGTAGGCTGTATCAGTTTCATTATGTGACTGGCCAGACCATGATGTTC 500
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
199 oSerSerIleAspProIleLeuGluLeuIleTrpAspValArgCysTyrG 216
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 TTCATCATTTGATTCTTCTGGACATGATAAGCTTAATGAGGAATATC 550
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
216 lnGluAspAspSerValProIleCysIleHisCysSerAlaGlyCysGly 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
551 AAGAACATGAGATGTTCTTATTGTTATTCATTCAGTCAGTCGANGCTGGA 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
233 ArgThrGlyValIleCysAlaIleValAspTyrThrTrpMetLeuLeuLy 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 AGAACAGTGCCATTTGTGCCATA...GATTATACGTGGGAATTTACTAAA 647
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 sAspGlyIleIleProGluAsnPheSerValPheSerLeuIleArgGluM 266
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
648 AGCTGGGAAAATCCAGAGGAATTTAATGTATTATTAATTAACAAGAAA 697
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
266 eArgThrGlnArgProSerLeuValGlnThrGlnGluGlnTyrGluLeu 282
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
698 TGAGAACACAAAGCATCTCTGAGTCNCAACAAAGGACCAATNTGAAC 747
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 ValTyrAsnAla 286
::: |||||
748 GGTCATAGAGCT 759

seq_name: gb_est2:BM469079
seq_documentation_block:
LOCUS BM469079 1003 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6455886 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587352
5', mRNA sequence.
ACCESSION BM469079
VERSION BM469079.1 GI:18518121
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1003)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12356 row: 1 column: 09
High quality sequence stop: 728.
FEATURES
location/Qualifiers
1..1003
/organism="Homo sapiens"
/db_xref="taxon:9606"

```



```

535 snProTyrPheSerSerTrpProSerGlyThrSerSerLysMetSer 551
|||||
100 ATCTTATTTTCATCATCGGCTCCAAAGTGTACCAGTTCTTAAGATGCT 149
|||||
552 LeuAspLeuProGluLysGlnAspGlyThrValPheProSerSerLeuLeu 568
|||||
150 CTTGATTTTACCTGAGAGCAAGATGGAACTGTTTTCCTTCTCTCTGTT 199
|||||
568 uprThrSerSerThrSerLeuPheSerTyrTyrAsnSerHisSerL 585
|||||
200 GCCAACATCCCTCTACATCCCTCTCTCTTATTAACAATTCACATGATCTT 249
|||||
585 euSerLeuAsnSerProThrAsnIleSerSerLeuLeuAsnGlnGluSer 601
|||||
250 TATCAGTGAATTCCTCAACCAATATTCCTCAGTATTGAACGAGAGTCA 299
|||||
602 AlavalLeuAlaThrAlaProArgIleaspaspGluIleProProLeu 618
|||||
300 GCTGTACTAGCAACTGCTCCAAAGATAGATGAATCCCCCTCCACT 349
|||||
618 uproValargThrProGluSerPheIleValValGluGluAlaGlyGluP 635
|||||
350 TCCGTGTCGGACACTGAATCATTTATGTGGTTGAGGAAGCTGGAGAA 399
|||||
635 heSerProAsnValProLysSerLeuSerSerAlaValLysValLysIle 651
|||||
400 TCTCACCAAATGTTCCCAATCCCTATCTCAGCTGTGAAGGTAATAAT 449
|||||
652 GlyThrSerLeuGluTrpGlyGlyThrSerGluProLysLysPheAspAs 668
|||||
450 GGAACATCAGTGGAAATGGGTGGAACTCTGAACCAAGAAATTTGATGA 499
|||||
668 pSerValIleLeuArgProSerLysSerValLysLeuArgSerPro 683
|||||
500 CTCGTGATGATAGACCAAGAGTGAATACTTCGAAAGTCTCT 545
|||||

```

seq_name: gb_est1:AW504119

seq_documentation_block:

LOCUS AW504119 475 bp mRNA linear EST 02-MAR-2000
DEFINITION UI-HF-BN0-ali-h-10-0-UI.r1 NIH_MGC_50 Homo sapiens cdna clone
IMAGE:3079915 5', mRNA sequence.

ACCESSION AW504119
VERSION AW504119.1 GI:7141786
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 475)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: M.B. Soares Lab

CDNA Library Arrayed by: M.B. Soares Lab

DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www.bio.lnl.gov/bbrp/image/image.html

Seq primer: MJ3 forward.

Location/Qualifiers

1. .475

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3079915"

/clone_lib="NIH_MGC_50"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

FEATURES

Source

BASE COUNT 134 a 87 c 115 g 139 t
ORIGIN
/lab_host="DH10B (LTI)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4Kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

alignment_scores:

Quality: 848.50 Length: 156
Ratio: 5.474 Gaps: 1
Percent Similarity: 99.359 Percent Identity: 99.359

alignment_block:

US-09-600-358A-4 x AW504119 ..

Align seg 1/1 to: AW504119 from: 1 to: 475

97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
|||||
10 AAGGCTTATATTGCCACCCAGGGTCTTTATCTACACCCCTCTGGACTT 59
|||||
113 eTrpArgMetIleTrpGluTyrSerValLeuIleIleValMetAlaCysM 130
|||||
60 CTGGAGGATGATTTGGGAATATAGTGTCTTATCATTTGTTGGCATGCA 109
|||||
130 etGluTyrGluMetGlyLysLysCysGluArgTyrTrpAlaGluPro 146
|||||
110 TGGAGTATGAAATGGGAAAGAAAAAGTGTAGCGCTACTGGGCTGAGCCA 159
|||||
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGl 163
|||||
160 GGAGAGATGCAGCTGGAATTTGGCCCTTCTCTGTATCTCTGTGAAGCTGA 209
|||||
163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsns 180
|||||
210 AAAAGGAAATCTGATTATATATAATCAGGACTCTAAAGCTTAAGTTCAATA 259
|||||
180 erGluThrArgThrThrIleTyrGlnPheHisTyrLysAsnTrpProAspHis 196
|||||
260 GTGAAACTCGAACTATCTACCAGTTTCATTACAAGATTGGCCAGACCAT 309
|||||
197 AspValProSerSerIleAspProIleLeuGluLeuIleTrpAspValAr 213
|||||
310 GATGTACCTTTCATCTATAGACCCCTATTCTTGAGCTCATCTGGGATGACG 359
|||||
213 qCysTyrGlnGluAspAspSerValProIleCysIleHisCysSerAlaG 230
|||||
360 TTGTTACCAAGAGGATGACAGTGTCCCATATGCAATTCACGACGTGCTG 409
|||||
230 lYcysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMet 246
|||||
410 GCTGTGGAAGGACTGGTGTATTGTTGCTATT...GATTATACATGGGATG 456
|||||
247 LeuLeuLysAspGlyIle 252
|||||
457 TTGCTAAAAGATGGGATA 474
|||||

seq_name: gb_est1:AUI32040

seq_documentation_block:

LOCUS AUI32040 811 bp mRNA linear EST 24-OCT-2000
DEFINITION AUI32040 NT2RP3 Homo sapiens cdna clone NT2RP303678 5', mRNA
sequence.

ACCESSION AUI32040

VERSION AUI32040.1 GI:10992394

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

451 ThrProPheGluLeuIleGlnArgGluThrLysGluValAspSerLy 467
 |||
 365 ACTCTTTTGATGATACACAGAGAAACCAAGGAGTGGACAGCA 414
 |||
 467 sGluAsnPheSerTyrLeuGluSerGlnProHisAspSerCysPheValG 484
 |||
 415 GGAAACCTTTCTTATTGGGAATCTCAACCCACATGATTCTTTGTTGTAG 464
 |||
 484 LuMetGlnAlaGlnLysValMetHisValSerSerAlaGluLeuAsnTyr 500
 |||
 465 AGATGCCGGCTCAACAAGTAATGCTGTTCTTCAACAGAACTGAATAT 514
 |||
 501 SerLeuProTyrAspSerLysGlnIleArgAsnAlaSerAsnValLy 517
 |||
 515 TCACCTGGCATATGACTCAACACAGGAATAGCTACTGCTCTACTGTAA 564
 |||
 517 sHisAspSerSerAlaLeu 524
 |||
 565 GCCCATGACTCTAGTGCTCTT 586

seq_name: gb_est1:AW915666

seq_documentation_block:
 LOCUS AW915666 554 bp mRNA linear EST 25-MAY-2000
 DEFINITION EST346970 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
 RGIC218 5' end, mRNA sequence.

ACCESSION AW915666
 VERSION AW915666.1 GI:8081371
 KEYWORDS EST.

SOURCE
 Rattus sp.

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 554)

AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,

Kerlavage, A.R. and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat

JOURNAL Gene Index

COMMENT Unpublished (1998)

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

This clone is available through the ATCC, contact the ATCC

tel#703-365-2700 for further information

Seq primer: M13 Reverse.

Location/Qualifiers

1..554

/organism="Rattus sp."

/db_xref="taxon:10118"

/clone="RGIC218"

/clone_lib="Normalized rat embryo, Bento Soares"

/dev_stage="embryo 8, 12, 18 dpc"

/note="vector: pT7T3pac; site_1: EcoRI; site_2: NotI"

BASE COUNT 166 a 123 c 129 g 136 t

ORIGIN

alignment_scores: Quality: 760.00 Length: 163

Ratio: 4.841 Gaps: 0

Percent Similarity: 96.319 Percent Identity: 86.503

alignment_block:

US-09-600-358A-4 x AW915666 ..

Align seg 1/1 to: AW915666 from: 1 to: 554

1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17

|||||
 66 ATGGACCAAGAGAAATTTCTGCAGCAGCTACTGAAAGAGGCCACAGAAA 115
 |||
 17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34
 |||
 116 GAAATTAACAGGGAGGAGTTTGCCCAATGAATTTCTGAAGCTGAAAGGC 165
 |||
 34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50
 |||
 166 AATCTACCAAGTACCAAGGCAGACAAAATATATCTTACAACCGTGGCTCAG 215
 |||
 51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67
 |||
 216 AGGCCCAAGAATATCAAGAAAAATAGATACAGGATATTTTGCCCTATGA 265
 |||
 67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSerT 84
 |||
 266 TCACAGCCTGGTAGAGTGCTCTGTTAACTGCGCATGAGGACTCCAGTT 315
 |||
 84 yrIleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100
 |||
 316 ACATCAATGCCAGCTTTATTAAAGGCTGCTATGGCCCGCCAGGCTTATATT 365
 |||
 101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetil 117
 |||
 366 GCCACTCAGGGTCTCTTATACAACTCTCTGGACTTCTTGGAGGATGAT 415
 |||
 117 eTrpGluTyrSerValLeuIleValMetAlaCysMetGluTyrGluM 134
 |||
 416 CTGGGAATACCGTGCTCTGGTCATCGTCATGCATGTCATGGAGTTCGAAA 465
 |||
 134 etGlyLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150
 |||
 466 TGGGAAAGAAAAGTGTGAGCGTTATTGGGCTGAGCCAGGAGACACAG 515
 |||
 151 LeuGluPheGlyProPheSerValSerCysGluAlaGlu 163
 |||
 516 CTGCAATTGGCCCTTTCTATATATCTCTGTGAAGACTGAG 554

seq_name: gb_est2:BG498752

seq_documentation_block:

LOCUS BG498752 745 bp mRNA linear EST 27-MAR-2001

DEFINITION 602544381F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:466653 5',

ACCESSION BG498752

VERSION BG498752.1 GI:13460269

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 745)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: CLONETECH Laboratories, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM1474 row: a column: 22

High quality sequence stop: 600.

FEATURES

source

1..745

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:466653"

/clone_lib="NIH_MGC_60"

```
/tissue_type="adenocarcinoma"  
/lab_host="DH10B (T1 phage-resistant)"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggccatagggc  
); Double-stranded cDNA was prepared from cell line RNA.  
5' and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CAGCGCATTTAGGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGGAGCGGCACATG-dT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC  
Library."
```

BASE COUNT 245 a 145 c 160 g 195 t
ORIGIN

alignment_scores:
Quality: 738.00 Length: 180
Ratio: 4.761 Gaps: 1
Percent Similarity: 86.111 Percent Identity: 85.000

alignment_block:
US-09-600-358A-4 x BG498752 ..
Align seg 1/1 to: BG498752 from: 1 to: 745

```
1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17  
|||||  
91 ATGGACCAAGAGAGAAATTCGCAAGAGTTCTCTGGATGAGGCCCAAGCAA 140  
17 sLysIleThrLysGluGluPheAlaAsnGluPheLeuLysLeuLysArg 34  
|||||  
141 GAAATTAATAAGAGAGAGTT.CCCATGAATTTCTGAAGCTGAAAGGC 189  
34 InSerThrLysTyrLysAlaAspLysThrTyrProThrThrValAlaGlu 50  
|||||  
190 AATCTACCAAGTACAGGAGAGCAAAACCTATCTACAACTGTGGCTGAG 239  
51 AsnAlaLysAsnIleLysLysAsnArgTyrLysAspIleLeuProTyrAs 67  
::: |||||  
240 AAGCCCAAGAATAATCAAGAAAAACAGATATAGGATATTTTGCCCTATGA 289  
67 pTyrSerArgValGluLeuSerLeuIleThrSerAspGluAspSerSert 84  
|||||  
290 TTATAGCCGGTAGAATAATCCCTGATAACCTCTGATGAGGATCCAGCT 339  
84 yTleAsnAlaAsnPheIleLysGlyValTyrGlyProLysAlaTyrIle 100  
|||||  
340 ACATCAATGCCAACTTCATTAGGGAGTTATGACCCCAAGGCTTATATT 389  
101 AlaThrGlnGlyProLeuSerThrThrLeuLeuAspPheTrpArgMetIl 117  
|||||  
390 GCCACCCAGGGTCTTATCTACAAACCTCTCTGGACTTCTGGAGGATGAT 439  
117 eTTPGluTyrSerValIleIleValMetAlaCysMetGluTyrClum 134  
|||||  
440 TTGGGAATATAGTGTCTTATCATCTGTTATGGCATGCATGGAGTATGAAA 489  
134 etGlyLysLysLysCysGluArgTyrTrpAlaGluProGlyGluMetGln 150  
|||||  
490 TGGGAAG..... 497  
151 LeuGluPheGlyProPheSerValSerCysGluAlaGluLysArgLysse 167  
|||||  
498 .....GAAGCTGAAAAAAGGAATC 517  
167 rAspTyrIleIleArgThrLeuLysValLysPheAsnSer 180  
|||||  
518 TGATTATATAATCAGGACTCTAAAGTTAAGTTCAATAGT 557
```

seq_name: gb_est2:BI152499

seq documentation block:
LOCUS BI152499 835 bp mRNA linear EST 05-JUL-2001
DEFINITION 602915640F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5066337 5',
mRNA sequence.
ACCESSION BI152499
VERSION BI152499.1 GI:14612500
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 835)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11178 row: o column: 10
High quality sequence stop: 762..

FEATURES

source Location/Qualifiers
1..835
/organism="Mus musculus"
/strain="C57BL/6J" (f1)
/db_xref="taxon:10090"
/clone="IMAGE:5066337"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary."
Stem cell origin.
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 261 a 151 c 210 g 213 t
ORIGIN

alignment_scores:
Quality: 718.50 Length: 304
Ratio: 3.421 Gaps: 7
Percent Similarity: 69.079 Percent Identity: 49.671

alignment_block:

US-09-600-358A-4 x BI152499 ..
Align seg 1/1 to: BI152499 from: 1 to: 835
1 MetAspGlnArgGluIleLeuGlnLysPheLeuAspGluAlaGlnSerLy 17
|||||
7 ATGAGCAAGTGGAGATCTCTGAGAGGTTCTATGGAGAGGTTCCAGGCCAT 56
17 sLysIleThr.....LysGluGluPheAlaAsnGluPheLeuL 30
|||||
57 GAAGAGTCCGACCAACAATGGGAGGAGCAACTCGCCCGGACATCATGT 105
30 ysLeuLysArgGlnSerThrLysTyrLysAlaAspLysThrTyrProThr 46
106 T..... 107
47 ThrValAlaGluAsnAlaLysAsnIleLysLysAsnArgTyrLysAspIl 63
107 107
63 eLeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspG 80
|||||

```
108 .....GATCACGCCGAGTTAGTGTGACTTTGAGATTCATCC 147
80 luAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyPro 96
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
148 RAGATTTCAGATTATATCAATGCAAAATTTATTAAAGGGTGTGTATGGCCA 197
97 LysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPh 113
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
198 AAAGCATATGTGGCAACCAAGGCGCTTGGGGAATACAGTCATAGACTT 247
113 eTrrpArgMetIleTrrpGluTyrSerValLeuIleIleValMetAlaCysM 130
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
248 CTGGAGAGTATGATGGAGTAAATGTTGTGATTATCGTGGCTGTC 297
130 eGluTyrGluMetGlyLysLysCysGluArgTyrTrrpAlaGluPro 146
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
298 GAGAAATTTGAGATGGGAGG..... 317
147 GlyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGl 163
   ||| ||
318 .....GAAATGA 325
163 uLysArgLysSerAspTyrIleIleArgThrLeuLysValLysPheAsnS 180
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
326 ACAAGCAAGACCGACTACTTCTCCGAACACTTTTACTTGAATTCAAA 375
180 eGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrrpProAspHis 196
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
376 ATGAATCCGTCGGCTCATCACTTTTCATGCTGAATGGCCAGACCAT 425
197 AspValProSerSerIleAspProIleLeuGluLeuIleTrrpAspVal 213
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
426 GATGTTCTCTCTGCTATTGATCTTATCTCGACATGATAAGCTTAATGAG 475
213 gCysTyrGlnGluAspAspSerValProIleCysIleHis.CysSerAla 229
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
476 GAAATACCAAGAACATGAGATGTCCTATTGTTATTCATTTGCAAGTGA 525
230 GlyCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrpMe 246
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
526 GGCTGTGGACGACAGGTGCTATTGTGCCATA...GATTACACGTGGAA 572
246 tLeuLeuLysAspGlyIleIleProGluAsnPheSerValPheSerLeuI 263
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
573 CTTACTGAAAGCAGGGAATTTCCAGAGAAATTAATGTATTAAATTTAA 622
263 leArgGluMetArgThrGlnArgProSerLeuValGlnThrGlnGluGln 279
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
623 TACAAGAAATGAGACACACAGAGGCACTCGGCAGTACAAACAAAGGAGCAG 672
280 TyrGluLeuValTyrAsnAlaValLeuGluLeuPheLysArgGlnMet 295
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
673 TATGAACTTGGTTATGAGAAGCTATGCTCAACTGTTTCCAAACAGGTA 722
296 AspVal 297
   :|||:
723 CAACGTG 728
```

seq_name: gb_est1:AW822729

seq_documentation_block:

LOCUS AW822729 629 bp mRNA linear EST 17-MAY-2000
DEFINITION uq17g09.y1 Ren Stubbs mouse thymus Mus musculus cDNA clone
IMAGE:2802784 5' similar to gb:X63440 M.musculus mRNA for
P19-protein tyrosine phosphatase (MOUSE);, mRNA sequence.

ACCESSION AW822729
VERSION AW822729.1 G1:7915806
KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 629)

AUTHORS

Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

TITLE

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1042308

Seq primer: Primer name ambiguous
High quality sequence stop: 464.

FEATURES

source

1..629
Location/Qualifiers
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:2802784"
/clone_lib="Ren Stubbs mouse thymus"
/sex="mixed"
/dev_stage="3 weeks"
/lab_host="DH10B"
/note="Organ: thymus; Vector: pT7T3D-Pac; Site_1: NotI;
Site_2: PacI; 1st strand cDNA was primed with an oligo(dT)
primer; double-stranded cDNA was ligated using 5' linker
ggcgctat and 3' linker aactgaagcttaatt. Library is
size-selected >2.5 kb and average insert size is 3.5 kb.
Clones were arrayed from primary plating; non-amplified.
Library constructed by X. Ren and L. Stubbs (Lawrence
Livermore National Laboratory and DOE Joint Genome
Institute, 7000 East Ave, L-453, Livermore, CA 94550)."

BASE COUNT 183 a 118 c 152 g 175 t 1 others
ORIGIN

alignment_scores:

Quality: 694.50 Length: 182
Ratio: 4.368 Gaps: 1
Percent Similarity: 87.363 Percent Identity: 67.582

alignment_block:

US-09-600-358A-4 x AW822729 ..

Align seg 1/1 to: AW822729 from: 1 to: 629

64 LeuProTyrAspTyrSerArgValGluLeuSerLeuIleThrSerAspGl 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
82 CTTCTGTTGATCAGACGCGAGTTAGTTGACTTTGAAGACTCCATCCCA 131
80 uAspSerSerTyrIleAsnAlaAsnPheIleLysGlyValTyrGlyProL 97
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 AGATTCAGATTATATCAATGCAAAATTTTATTAAAGGCTGTATGGGCCAA 181
97 ysAlaTyrIleAlaThrGlnGlyProLeuSerThrThrLeuLeuAspPhe 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
182 AAGCATATGTGGCAACCCAGGCGCTTTGGCGAATACAGTCATAGACTTC 231
114 TrrpArgMetIleTrrpGluTyrSerValLeuIleIleValMetAlaCysMe 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 TGGAGATGATATGGGAGTATATGTTGTATCATCTGATGCCCTTGTGTCG 281
130 tGluTyrGluMetGlyLysLysCysGluArgTyrTrrpAlaGluProG 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
282 AGAATTTGAGATGGGAAGAAAAAGTGTGAGCGCTACTGCGCTTTGTATG 331
147 lyGluMetGlnLeuGluPheGlyProPheSerValSerCysGluAlaGlu 163
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

332 GAGNAGATCCTATAAACATTTCACACCATTAATAATTTCTGTGAAANTGAA 381

164 LysArgLysSerAspTyrIleileArgThrLeuLysValLysPheAsnSe 180
::: :::::GAGGACCGGACTACTTCATCCGACACATTTTACATTTGAATTCAAAA 431
382 CAAGCAAGGACCGGACTACTTCATCCGACACATTTTACATTTGAATTCAAAA 431

180 rGluThrArgThrIleTyrGlnPheHisTyrLysAsnTrpProAspHisA 197
::: :::::GAGGACCGGACTACTTCATCCGACACATTTTACATTTGAATTCAAAA 431
432 TGAATCCCGTCGGCTCATACAGTTTCATTAGTGAACCTGGCCAGACCATG 481

197 spValProSerSerIleaspProfileLeuGluLeuIleTyrAspValArg 213
::: :::::GAGGACCGGACTACTTCATCCGACACATTTTACATTTGAATTCAAAA 431
482 ATGTTCCTTCGTCATTGATTCATTCATTCGACATGATAGCTTAATGAGG 531

214 CysTyrGlnGluAspSerValprofileCysIleHisCysSerAlaAcI 230
::: :::::GAGGACCGGACTACTTCATCCGACACATTTTACATTTGAATTCAAAA 431
532 AAATACCAGACATGAAGATGTCCTCATATGTCATTCATTCAGATGCAGG 581

230 yCysGlyArgThrGlyValIleCysAlaIleValAspTyrThrTrp 245

582 CTGTGGACGAACAGGTGCTATTTCTGGCATA...GATTACACGTGG 624

seq_name: qb_est2:BM194577

seq_documentation block:

LOCUS	BM194577	505 bp	mRNA	linear	EST 13-DEC-2001
DEFINITION	TCAAP3E10089 Pediatric pre-B cell acute lymphoblastic leukemia Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCAAP1008, mRNA sequence.				

Accession BM194577
Version BM194577.1
Keywords FST

SOURCE	ORGANISM	Homo sapiens

Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 505)

AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr., Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

TITLE Pediatric Leukemia cDNA Sequencing Project (2001)

JOURNAL
UNPUBLISHED (2001)
UNPUBLISHED (2001)

COMMENT
Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA

FEATURES

source

```

1: 300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCAAP1008"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGGAGACTCGAGCGCGCAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'GAGAGTCGAGTCGCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with
BamHI and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2002, 08:00:48 ; Search time 474.64 seconds
(without alignments)
8522.355 Million cell updates/sec

Title: US-09-600-358a-3
Perfect score: 2356

Sequence: 1 tccctcaacctacttataga.....cctagtcggagcacatccgg 2356

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2356	100.0	2356	20	Human Cytoplasmic
2	2097	89.0	3058	20	Human Cytoplasmic
3	2074.6	88.1	3580	20	Human PTP04 encodi
4	1729	73.4	2452	22	Human tyrosine pho
5	621.4	26.4	1286	21	Human breast and o
6	368	15.6	2983	24	Mouse ischaemic co
7	216.4	9.2	1529	18	cDNA encoding a no
8	214.2	9.1	2226	19	Rat protein tyrosi
9	202.2	8.6	2810	19	Human brain derive

10	160	6.8	205	21	AA12676	Human secreted pro
11	138.8	5.9	4798	22	AA158418	Human polynucleoti
12	130.6	5.5	2309	15	AA073786	Partial PCR fragme
13	129	5.5	4965	22	AAH98348	Human EST-derived
14	129	5.5	4968	22	AA160204	Human polynucleoti
15	120	5.1	466	18	AAV03128	Partial cDNA encod
16	116.8	5.0	5117	16	AAV06027	huDEP-1 cDNA. Hom
17	115.4	4.9	3969	19	AAV05762	Chicken protein ty
18	115.4	4.9	3969	19	AAV81897	Chicken transmembr
19	115.2	4.9	2674	22	AAH99616	Human protein enco
20	115.2	4.9	5069	22	AAK52292	Human polynucleoti
21	115.2	4.9	5164	22	AAK53276	Human polynucleoti
22	115.2	4.9	5465	22	AAK50778	Human cancer agent
23	112.8	4.8	4555	16	AAQ94311	Tyrosine phosphata
24	106.2	4.5	247	18	AAV69301	Murine metastatic
25	104.8	4.4	4752	19	AAV04868	Nucleotide sequenc
26	104.8	4.4	4799	19	AAV04887	Nucleotide sequenc
27	104.8	4.4	4815	19	AAV04886	Nucleotide sequenc
28	104.8	4.4	5113	17	AAV48479	Human tyrosine dep
29	104	4.4	2760	24	AB199784	Mouse ischaemic co
30	100.6	4.3	5455	19	AAV33577	Osteosteticular pr
31	98.4	4.2	4651	15	AAQ72913	Human receptor-tyr
32	98.4	4.2	4823	23	AAK52612	DNA encoding novel
33	98.4	4.2	5775	20	AAZ08539	Human protein phos
34	92.4	3.9	6075	21	AAK88866	Human protein tyro
35	92.4	3.9	6342	22	AAK51954	Human polynucleoti
36	92.4	3.9	6346	22	AAK52938	Human polynucleoti
37	92.4	3.9	6632	22	AAZ26742	Human breast cance
38	90.4	3.8	3623	20	AAV82821	Rabbit osteoclast-
39	90.4	3.8	4573	22	AAH98379	Human EST-derived
40	90.4	3.8	5412	24	AB199920	Rat mucocardial ce
41	90.4	3.8	5690	16	AAQ86902	Rat receptor type-
42	90	3.8	6000	16	AAQ86478	Human PTP-OB. Hom
43	90	3.8	6000	18	AAK85389	Human protein tyro
44	90	3.8	6000	20	AAK06095	Human protein tyro
45	90	3.8	6000	22	AAK02144	Human protein tyro

ALIGNMENTS

RESULT 1

AAK90696

ID AAK90696 standard; cDNA; 2356 BP.

AC AAK90696;

XX

XX 01-OCT-1999 (first entry)

DE Human Cytoplasmic phosphatase Lyp2, cDNA.

XX Lymphoid Protein Tyrosine Phosphatase gene; Lyp gene; immunosuppressant;
KW intracellular tyrosine phosphatase; PTPase; lymphocyte; fetal liver;
KW resting lymphoid cell; protein tyrosine kinase; PKs; transgenic animal;
KW lymphoma; T cell antigen receptor signalling; intronic sequence;
KW cytokine receptor signalling; protein tyrosine phosphatase; PTPase; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 42..2120

FT /*tag= a

FT /product= "Cytoplasmic phosphatase, Lyp2"

FT /note= "Expressed in fetal liver tissue and resting lymphoid cells"

FT WO9936548-A1.

PN

XX 22-JUL-1999.

PD

XX 18-JAN-1999; 99WO-CA00038.

XX 16-JAN-1998; 98CA-2220853.

PR

XX (HSCR-) HSC RES & DEV LP.
PA Roifman CM;
XX
PI
XX
DR WPI; 1999-444404/37.
XX P-PSDB; AAY28653.
DR
XX
PT New nucleic acid encoding intracellular tyrosine phosphatase and
PT related proteins, used to modulate signalling through T cells,
PT particularly as immunosuppressant
XX
PS Claim 5b; Page 54; 105pp; English.
XX
CC The present cDNA sequence is that of the Lymphoid protein Tyrosine
CC Phosphatase gene, Lyp2 that encodes a cytoplasmic tyrosine phosphatase
CC protein with a single catalytic domain. The Lyp gene has been localised
CC to human chromosome lp13. It is expressed significantly in fetal liver
CC and in resting lymphoid cells. The Lyp2 gene sequence arises by
CC alternative splicing of the Lyp1 mRNA and the coding sequence reads into
CC the intronic sequence of Lyp1 till a termination codon is encountered.
CC Lyp gene products are important for regulation of T cell antigen and
CC cytokine receptor signalling and for early and late stages of T cell
CC differentiation. Lyp2 has immunosuppressive activity. This sequence is
CC used to treat lymphoma, to reduce or prevent T cell activation or
CC proliferation and to control thymocyte differentiation. Fragments of
CC this gene are used, as probes or primers, to identify allelic variations
CC and for diagnosis of diseases associated with reduced Lyp activity.
CC Transgenic animals in which a Lyp gene has been inactivated or replaced
CC by a heterologous Lyp gene are used to study and screen for possible
CC therapeutic agents.
XX
SQ Sequence 2356 BP; 750 A; 466 C; 466 G; 674 T; 0 other;

Query Match 100.0%; Score 2356; DB 20; Length 2356;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcctcaactctatagactattttctgctgcagcatggaccacaaagagaattc 60
DB 1 tcctcaactctatagactattttctgctgcagcatggaccacaaagagaattc 60
QY 61 tgcagaagttcctggatgagcccaagcaagaaaattactaaaggagttgccaatg 120
DB 61 tgcagaagttcctggatgagcccaagcaagaaaattactaaaggagttgccaatg 120
QY 121 aattctgaagctgaaaaaggaatctaccaagtacaaaggcagacaaaacctatcctacaa 180
DB 121 aattctgaagctgaaaaaggaatctaccaagtacaaaggcagacaaaacctatcctacaa 180
QY 181 ctgtgctgaagaatgccaaagaatatacaaaaaaacagataagatattttgcctatg 240
DB 181 ctgtgctgaagaatgccaaagaatatacaaaaaaacagataagatattttgcctatg 240
QY 241 attatagccgggtagaaactatccctgataacctctgatgaggattccagctacatcaatg 300
DB 241 attatagccgggtagaaactatccctgataacctctgatgaggattccagctacatcaatg 300
QY 301 ccaacttcattaaaggagtttatggacccaaggcttatattgcccacccgggtccctttat 360
DB 301 ccaacttcattaaaggagtttatggacccaaggcttatattgcccacccgggtccctttat 360
QY 361 ctacaacctctcgactctgagatgatttgggaataatagtgcttcttatcatgttta 420
DB 361 ctacaacctctcgactctgagatgatttgggaataatagtgcttcttatcatgttta 420
QY 421 tggcatgcatggagtgataaagggaagaaaagtgtagcgctactggtggtgagccag 480
DB 421 tggcatgcatggagtgataaagggaagaaaagtgtagcgctactggtggtgagccag 480
QY 481 gagagatgcagctggaaatttggcccttctctgtatcctctgtagctgaaaaaagaaat 540
DB 481 gagagatgcagctggaaatttggcccttctctgtatcctctgtagctgaaaaaagaaat 540

DB 481 gagagatgcagctggaaatttggcccttctctgtatcctctgtagctgaaaaaagaaat 540
QY 541 ctgattatataatcagagactctaaaagttaagttaacttaagttaacttaactatctacc 600
DB 541 ctgattatataatcagagactctaaaagttaagttaacttaagttaacttaactatctacc 600
QY 601 agttcattacaagaatttgccagaccatgatgaccttcacatctatagaccctattcttg 660
DB 601 agttcattacaagaatttgccagaccatgatgaccttcacatctatagaccctattcttg 660
QY 661 agctcatctgggatgtacgtttgtaccagagagatgacaggtgtcccatatgcatcaact 720
DB 661 agctcatctgggatgtacgtttgtaccagagagatgacaggtgtcccatatgcatcaact 720
QY 721 gaagtctggtgtgaaaggactggttatttgttctgtatttgttattatatacatgatgt 780
DB 721 gaagtctggtgtgaaaggactggttatttgttctgtatttgttattatatacatgatgt 780
QY 781 tgcataaagatggataaattcctcagaaactcagtttctcagtttgcgttcgggaaatgc 840
DB 781 tgcataaagatggataaattcctcagaaactcagtttctcagtttgcgttcgggaaatgc 840
QY 841 ggacacagagcccttcatttagttcacaacgaggaacaatatgaactggtctacaatgctg 900
DB 841 ggacacagagcccttcatttagttcacaacgaggaacaatatgaactggtctacaatgctg 900
QY 901 tattagaactatttaagacagatgatttatcagagataaaacattctggacagaga 960
DB 901 tattagaactatttaagacagatgatttatcagagataaaacattctggacagaga 960
QY 961 gtcaagcaaaagcattgtattctcctgagaaaaatcacactctccagcagactcttattctc 1020
DB 961 gtcaagcaaaagcattgtattctcctgagaaaaatcacactctccagcagactcttattctc 1020
QY 1021 ctaatttacaaaaagtaccacaaaagcgaacaaaatgatgaaccacaaagacacaaaa 1080
DB 1021 ctaatttacaaaaagtaccacaaaagcgaacaaaatgatgaaccacaaagacacaaaa 1080
QY 1081 tggaaatcaaaagaattctctctcttgacttttaggactcttgaaataagtgcacaaagaag 1140
DB 1081 tggaaatcaaaagaattctctctcttgacttttaggactcttgaaataagtgcacaaagaag 1140
QY 1141 agctagttttgacacctgctaaatcaagcactcttttgcactttctggagctaaattaca 1200
DB 1141 agctagttttgacacctgctaaatcaagcactcttttgcactttctggagctaaattaca 1200
QY 1201 gttttgacaaaaatgctgacacaaacctgaaatggcagacaaaggcatttcccaatggtg 1260
DB 1201 gttttgacaaaaatgctgacacaaacctgaaatggcagacaaaggcatttcccaatggtg 1260
QY 1261 gggagcctcttcagaagcatcaaaagtttgatttgggctctctttgttggaggatggt 1320
DB 1261 gggagcctcttcagaagcatcaaaagtttgatttgggctctctttgttggaggatggt 1320
QY 1321 ctaattctaaacctgaaatgcagcaggaagatatttttaattcaaaaggtgccaaatacac 1380
DB 1321 ctaattctaaacctgaaatgcagcaggaagatatttttaattcaaaaggtgccaaatacac 1380
QY 1381 ggacccaaatacaactccttttgaattgatacagcagagagaacaaaggaggtggacagca 1440
DB 1381 ggacccaaatacaactccttttgaattgatacagcagagagaacaaaggaggtggacagca 1440
QY 1441 aggaacacttttctatttgggaatctcaaccacatgatctcttcttctgtagagatgcagg 1500
DB 1441 aggaacacttttctatttgggaatctcaaccacatgatctcttcttctgtagagatgcagg 1500
QY 1501 ctcaaaaagtaatgcattgtttcttcagcagaactgaattattcactgccaatactactcta 1560
DB 1501 ctcaaaaagtaatgcattgtttcttcagcagaactgaattattcactgccaatactactcta 1560
QY 1561 aacacccaaatacgttaatgcctctaaatgaagacacacatgacttagtgcctctgggtgat 1620
DB 1561 aacacccaaatacgttaatgcctctaaatgaagacacacatgacttagtgcctctgggtgat 1620


```
QY 361 ctacaacctcctggactcttgagatgatttgggaataatagtgctcttatctgtta 420
|||
Db 361 ctacaacctcctggactcttgagatgatttgggaataatagtgctcttatctgtta 420
|||
QY 421 tggcatgcatggaatggaatgggaagaaagtgtagcgctactggcgtagccag 480
|||
Db 421 tggcatgcatggaatggaatgggaagaaagtgtagcgctactggcgtagccag 480
|||
QY 481 gagagatgagctggaatttggcccttctctatctcttgtagagtgaaaaagaaat 540
|||
Db 481 gagagatgagctggaatttggcccttctctatctcttgtagagtgaaaaagaaat 540
|||
QY 541 ctgattataaactcaggactctaaaagttaagtccaatagtagaaactcgaaactatacc 600
|||
Db 541 ctgattataaactcaggactctaaaagttaagtccaatagtagaaactcgaaactatacc 600
|||
QY 601 agttcattacaagaattggccagaccatgtagtacccttcattatagaccctattcttg 660
|||
Db 601 agttcattacaagaattggccagaccatgtagtacccttcattatagaccctattcttg 660
|||
QY 661 agctcattcggtatgacgttcttacaagagagatgacagtggtcccatatgattcaact 720
|||
Db 661 agctcattcggtatgacgttcttacaagagagatgacagtggtcccatatgattcaact 720
|||
QY 721 gcagtgctggctgtggaagactggtgtatttggctatttggattatcacatgagatgt 780
|||
Db 721 gcagtgctggctgtggaagactggtgtatttggctatttggattatcacatgagatgt 780
|||
QY 781 tgcataaagatggataattcctggaacttcagtggttttcagtttgatccgggaatgc 840
|||
Db 781 tgcataaagatggataattcctggaacttcagtggttttcagtttgatccgggaatgc 840
|||
QY 841 ggcacagagcctcatttagttcaaacgcaggaacaatatgaactggtctacaatgctg 900
|||
Db 841 ggcacagagcctcatttagttcaaacgcaggaacaatatgaactggtctacaatgctg 900
|||
QY 901 tattagaactatttaagagacagatggatgttatcagagataaacattctgggaacagaga 960
|||
Db 901 tattagaactatttaagagacagatggatgttatcagagataaacattctgggaacagaga 960
|||
QY 961 gtaagcaaaagcattgtattcctgagaaaaatcacactctccaagcagactcttattctc 1020
|||
Db 961 gtaagcaaaagcattgtattcctgagaaaaatcacactctccaagcagactcttattctc 1020
|||
QY 1021 ctaatttaccaaaaagtaccacaaaagcagcaaaaaatgatgaaccacaaagacaaaaa 1080
|||
Db 1021 ctaatttaccaaaaagtaccacaaaagcagcaaaaaatgatgaaccacaaagacaaaaa 1080
|||
QY 1081 tggaaatcaaagaactctctcttggacttttaggactcttgaaataaagtgcataaagaag 1140
|||
Db 1081 tggaaatcaaagaactctctcttggacttttaggactcttgaaataaagtgcataaagaag 1140
|||
QY 1141 agctagtttgcacctgctaaatacagaactctcttggactttctggagactaaattaca 1200
|||
Db 1141 agctagtttgcacctgctaaatacagaactctcttggactttctggagactaaattaca 1200
|||
QY 1201 gttttgacaaaaatgctgcacacacacatgaaatggcagacaaaaggcatttcccaatgttg 1260
|||
Db 1201 gttttgacaaaaatgctgcacacacacatgaaatggcagacaaaaggcatttcccaatgttg 1260
|||
QY 1261 gggagcctcttcagaagcatcaaatgttgatttgggctctcttggtaggggagttt 1320
|||
Db 1261 gggagcctcttcagaagcatcaaatgttgatttgggctctcttggtaggggagttt 1320
|||
QY 1321 ctaattctaaacctgtaaatgcagcaggagaagatttttaattcaaaaggtgccataaacac 1380
|||
Db 1321 ctaattctaaacctgtaaatgcagcaggagaagatttttaattcaaaaggtgccataaacac 1380
|||
QY 1381 ggcacaaatacaactccttttgaattgatacagcagagagaaacaaaggaggtggacagca 1440
|||
Db 1381 ggcacaaatacaactccttttgaattgatacagcagagagaaacaaaggaggtggacagca 1440
|||
QY 1441 aggaaaaactttcttatttggaaatctcaaccacatgattctgtttttagagatgcagg 1500
|||
```

```
|||
Db 1441 aggaaaaactttcttatttggaaatctcaaccacatgattctgtttttagagatgcagg 1500
|||
QY 1501 ctcaaaaaagtaatgcatgtgtttcttcagcagaactgaattattcaactgcatatgactcta 1560
|||
Db 1501 ctcaaaaaagtaatgcatgtgtttcttcagcagaactgaattattcaactgcatatgactcta 1560
|||
QY 1561 acaccaaataagtaatgcatgtgttaagttaagcaccatgactctagtctcttgggtgtat 1620
|||
Db 1561 acaccaaataagtaatgcatgtgttaagttaagcaccatgactctagtctcttgggtgtat 1620
|||
QY 1621 attcttcacatcattttagtggaaaatccttatttttcatcattggcctccaagtggtaacca 1680
|||
Db 1621 attcttcacatcattttagtggaaaatccttatttttcatcattggcctccaagtggtaacca 1680
|||
QY 1681 gttctaaagtgtctctgtatttacctgagaagcaaatggaactgttttctctctctc 1740
|||
Db 1681 gttctaaagtgtctctgtatttacctgagaagcaaatggaactgttttctctctctc 1740
|||
QY 1741 tgtgtcacaacatctctacatccctctctcttatttacaattcacatagttctttatcac 1800
|||
Db 1741 tgtgtcacaacatctctacatccctctctcttatttacaattcacatagttctttatcac 1800
|||
QY 1801 tgaattctccaaaccaatatttctcactatttgaaccaggagtcagctgtactagaacctg 1860
|||
Db 1801 tgaattctccaaaccaatatttctcactatttgaaccaggagtcagctgtactagaacctg 1860
|||
QY 1861 ctccaaagtagatgatgaaatccccctccactctctctctgaaggacacacacgaatcattta 1920
|||
Db 1861 ctccaaagtagatgatgaaatccccctccactctctctctgaaggacacacacgaatcattta 1920
|||
QY 1921 ttgtgtgtgaggaagctggagaattctcccaaatgttcccaaatctctatctcagctg 1980
|||
Db 1921 ttgtgtgtgaggaagctggagaattctcccaaatgttcccaaatctctatctcagctg 1980
|||
QY 1981 tgaaggtaaaaatttgaacatcacactggaatgggttggaaacatctgaaccacaaagaaatttg 2040
|||
Db 1981 tgaaggtaaaaatttgaacatcacactggaatgggttggaaacatctgaaccacaaagaaatttg 2040
|||
QY 2041 atgactctgtgatacttagaccaagaagagtgtaaaactccgaagctctaaatcacg 2097
|||
Db 2041 atgactctgtgatacttagaccaagaagagtgtaaaactccgaagctctaaatcacg 2097
|||
RESULT 3
AAV81742
ID AAV81742 standard; cDNA; 3580 BP.
XX
AC AAV81742;
XX
DT 10-MAR-1999 (first entry)
XX
DE Human PTP04 encoding cDNA.
XX
KW PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;
type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;
neurodegenerative disease; neuronal survival; Alzheimer's disease;
Parkinson's disease; Huntington's disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 53..2476
FT /*tag= a
XX
PN W09849317-A2.
PD 05-NOV-1998.
XX
PF 27-APR-1998; 98WO-US08439.
XX
PR 23-OCT-1997; 97US-0063595.
28-APR-1997; 97US-0044428.
PR
```

PR 20-MAY-1997; 97US-0047222.
PR 11-JUN-1997; 97US-0049477.
PR 11-JUN-1997; 97US-0049756.
PR 18-JUN-1997; 97US-0049914.
XX (SUGE-) SUGEN INC.
FA
XX App H, Clary D, Courtneidge SA, Hui TH, Jallal B;
PI Markby D, Onrust S, Peles E, Plowman GD;
XX
DR WPI: 1999-009434/01.
DR P-PSDB: AAW89247.
XX
XX New nucleic acid encoding specific protein tyrosine phosphatases -
PT useful for identifying specific modulators for treatment and
PT prevention of cancer and neurodegenerative disease
XX
PS Claim 2; Page 145-146; 193pp; English.
XX
CC The present invention describes isolated, enriched or purified nucleic
CC acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The
CC present sequence encodes human PTP04. The above proteins, other than
CC ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify
CC substances that modulate their activity (i.e. agonists and antagonists,
CC including NBP) in vivo or in vitro. These substances are used to treat
CC or prevent diseases associated with abnormal signal transduction
CC pathways that involve the proteins, particularly cancer (e.g. leukemia
CC and lymphoma), while modulators of ALK-7 (which is a type 1 receptor
CC serine/threonine kinase) are used to promote neuronal survival,
CC particularly for treating Alzheimer's, Parkinson's or Huntington's
CC diseases. Nucleic acid fragments of the polynucleotides encoding the
CC proteins can be used as probes to identify and clone related sequences;
CC to detect protein-encoded RNA; to generate transgenic animals and in
CC gene therapy (optionally after mutation). Ab are used to determine the
CC proteins.
XX
SQ Sequence 3580 BP; 1213 A; 666 C; 618 G; 1083 T; 0 other;

Query Match 88.18; Score 2074.6; DB 20; Length 3580;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2090; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 1 tcctccaacctactatagactattttctgctcgcagcatggaccacaaagagaaattc 60
Db 12 tcctccaacctactatagactattttctgctcgcagcatggaccacaaagagaaattc 71
Qy 61 tgcagaagttcctgagtgagcccaaaagcaagaaattactaaagagaggttgcgaatg 120
Db 72 tgcagaagttcctgagtgagcccaaaagcaagaaattactaaagagaggttgcgaatg 131
Qy 121 aatttgaagctgaaaggaattctaccagtaacaaagtaacaaagcaaaacctatctacaa 180
Db 132 aatttgaagctgaaaggaattctaccagtaacaaagtaacaaagcaaaacctatctacaa 191
Qy 181 ctgtggctgagaatgccaaagatatcaagaaacagagataaaggtatttgcctcatg 240
Db 192 ctgtggctgagaagcccaagatatcaagaaacagagataaaggtatttgcctcatg 251
Qy 241 attatgacgggtgagaactatccctgataaactctgtgaggtattccagctacatcaatg 300
Db 252 attatgacgggtgagaactatccctgataaactctgtgaggtattccagctacatcaatg 311
Qy 301 ccaacttcaataaaggagttatggacccaaggcttatattgccaacccagggtcccttat 360
Db 312 ccaacttcaataaaggagttatggacccaaggcttatattgccaacccagggtcccttat 371
Qy 361 ctacaacctctgactctcggagatgatttgggaataagtgctcttatcatgttta 420
Db 372 ctacaacctctgactctcggagatgatttgggaataagtgctcttatcatgttta 431
Qy 421 tggcatgcatgagtatgaaatggaaagaaagaaagtgtgacgctactgggctgagccag 480
Db 1213 A; 666 C; 618 G; 1083 T; 0 other;

Db 432 tggcatgcatgagtatgaaatggaaagaaagtgagcgctactgggctgagccag 491
Qy 481 gagagatcagctgaaatttggcccttctctgtatctctgtgaagctgaaagaaagaaat 540
Db 492 gagagatcagctgaaatttggcccttctctgtatctctgtgaagctgaaagaaagaaat 551
Qy 541 ctgattataatacagagactctaaaagttaagttaagttaagttaagttaagttaagttaagt 600
Db 552 ctgattataatacagagactctaaaagttaagttaagttaagttaagttaagttaagttaagt 611
Qy 601 agtttctatacaagaatttggccagaccatgatgtaccctctatctatagaccctattctgt 660
Db 612 agtttctatacaagaatttggccagaccatgatgtaccctctatctatagaccctattctgt 671
Qy 661 agtcatctggatgacgttcttaccagagagatgacagtggttcccatgacattcact 720
Db 672 agtcatctggatgacgttcttaccagagagatgacagtggttcccatgacattcact 731
Qy 721 gcagtgcgtggttggaggactggttatttgcgtatttgcgtatttgcgtatttgcgtatttgc 780
Db 732 gcagtgcgtggttggaggactggttatttgcgtatttgcgtatttgcgtatttgcgtatttgc 788
Qy 781 tgcataaagatgggataattcctcagaaactcgaactcgaactcgaactcgaactcgaactcga 840
Db 789 tgcataaagatgggataattcctcagaaactcgaactcgaactcgaactcgaactcgaactcga 848
Qy 841 ggacacagagcccttcattagttcaaacgcaggaacaatatgaactggtctacaaatgctg 900
Db 849 ggacacagagcccttcattagttcaaacgcaggaacaatatgaactggtctacaaatgctg 908
Qy 901 tattagaactatttaagagacagatggatgttatcagagataaaacattctcgtgaacagaga 960
Db 909 tattagaactatttaagagacagatggatgttatcagagataaaacattctcgtgaacagaga 968
Qy 961 gtcaagcaaaacattgtattcctcagaaataacacactcctccaaagcagactctattctc 1020
Db 969 gtcaagcaaaacattgtattcctcagaaataacacactcctccaaagcagactctattctc 1028
Qy 1021 ctaatttaccaaaagtaaccacaaaagcagcaaaaatgatgaacaaacaaagagcaaaaa 1080
Db 1029 ctaatttaccaaaagtaaccacaaaagcagcaaaaatgatgaacaaacaaagagcaaaaa 1088
Qy 1081 tggaaatcaaaagaattcttctctcttgccttgccttgccttgccttgccttgccttgccttgc 1140
Db 1089 tggaaatcaaaagaattcttctctcttgccttgccttgccttgccttgccttgccttgccttgc 1148
Qy 1141 agctagtttgcacctgctaaatcaaacactctcttgccttgccttgccttgccttgccttgccttgc 1200
Db 1149 agctagtttgcacctgctaaatcaaacactctcttgccttgccttgccttgccttgccttgccttgc 1208
Qy 1201 gttttgacaaaatgctgacacaaacctgaaatggcagacaaagcatttccaatagttg 1260
Db 1209 gttttgacaaaatgctgacacaaacctgaaatggcagacaaagcatttccaatagttg 1268
Qy 1261 gggagcctcttcagaagcatcaaaagtgttgattgggtctctcttcttcttcttcttcttcttcttct 1320
Db 1269 gggagcctcttcagaagcatcaaaagtgttgattgggtctctcttcttcttcttcttcttcttcttct 1328
Qy 1321 ctaattcacaacctgtaaatgcagcaggaagataatttcaattcaaaaggtgcacaataaacac 1380
Db 1329 ctaattcacaacctgtaaatgcagcaggaagataatttcaattcaaaaggtgcacaataaacac 1388
Qy 1381 ggacaaaatcaacctcttcttgaattgatcacagacagaaacaaagaggtggagacaga 1440
Db 1389 ggacaaaatcaacctcttcttgaattgatcacagacagaaacaaagaggtggagacaga 1448
Qy 1441 agggaaactcttcttcttgaattgatcac 1500
Db 1449 agggaaactcttcttcttgaattgatcac 1508
Qy 1501 ctcaaaaagtaagtgttcttctcagcagaactgaaattattcactgccaatgactacta 1560
Db 1509 ctcaaaaagtaagtgttcttctcagcagaactgaaattattcactgccaatgactacta 1568

QY 1561 aacacaaaatagccttaataatgtaagcaccatgactctagtctcttggtgtat 1620
|||||
Db 1569 aacacaaaatagccttaataatgtaagcaccatgactctagtctcttggtgtat 1628
|||||
QY 1621 attctacatccttttagtggaaaatcccttatttttctcatcctggcctccaagtgggtacca 1680
|||||
Db 1629 attctacatccttttagtggaaaatcccttatttttctcatcctggcctccaagtgggtacca 1688
|||||
QY 1681 gttctaagatgtctcttgatttaccctgagaagcaaatggaactgtttttctctcttc 1740
|||||
Db 1689 gttctaagatgtctcttgatttaccctgagaagcaaatggaactgtttttctctcttc 1748
|||||
QY 1741 tgttgccaacatcctctacatccctctctcttatttaccattcacatagttcttttatcac 1800
|||||
Db 1749 tgttgccaacatcctctacatccctctctcttatttaccattcacatgattcttttatcac 1808
|||||
QY 1801 tgaattctccaacccaatatttctcactattgaaccagagtcagctgtactagcaactg 1860
|||||
Db 1809 tgaattctccaacccaatatttctcactattgaaccagagtcagctgtactagcaactg 1868
|||||
QY 1861 ctccaagatagatgatgaataccccctccacttctctgtacggacacactggaatcattta 1920
|||||
Db 1869 ctccaagatagatgatgaataccccctccacttctctgtacggacacactggaatcattta 1928
|||||
QY 1921 ttgtggttgagggaagctggagaattctccacaaaatgttcccaaatcttctatccctcagctg 1980
|||||
Db 1929 ttgtggttgagggaagctggagaattctccacaaaatgttcccaaatcttctatccctcagctg 1988
|||||
QY 1981 tgaagtgaaaaattgaaacatcactggaatggggtggaacatctctgaacaaaagaaatttg 2040
|||||
Db 1989 tgaagtgaaaaattgaaacatcactggaatggggtggaacatctctgaacaaaagaaatttg 2048
|||||
QY 2041 atgactctgtactatagaccagcaagagtgtaaaactccgaactccgaacttaaatcag 2097
|||||
Db 2049 atgactctgtactatagaccagcaagagtgtaaaactccgaactccgaacttaaatcag 2105
|||||

RESULT 4
AAH79342
ID AAH79342 standard; DNA; 2452 BP.
XX
XX AAH79342;
XX AC
XX XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase coding sequence.
XX
XX Human; tyrosine phosphatase; hPTP; ds.
XX
XX Homo sapiens.
XX OS
XX CN1302899-A.
XX
XX 11-JUL-2001.
XX PD
XX 29-OCT-1999; 99CN-0119935.
XX PF
XX 29-OCT-1999; 99CN-0119935.
XX PR
XX
XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX PA
XX Ren S, Wu T, Qian B;
XX PI
XX WPI; 2001-550591/62.
XX DR P-PSDB; AAG78623.
XX
XX Human protein tyrosine phosphatase and its coding sequence -
XX
XX Claim 1; Page 19-21(Disclosure); 29pp; Chinese.
XX PS
XX The present invention provides the protein and coding sequences of human
XX CC tyrosine phosphatase hPTP. The protein is expressed in human normal
CC

CC suprarenal tissue. The present sequence is the coding sequence of the
XX invention.
SQ Sequence 2452 BP; 791 A; 516 C; 468 G; 677 T; 0 other;

Query Match 73.4%; Score 1729; DB 22; Length 2452;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 100; Indels 27; Gaps 13;

QY 33 ctctgcagcatggaccacaaagaaaattctgcagaagtctctctggatgagggcccaaaagcaag 92
|||||
Db 1 ctctgcagcatggaccacaaagaaaattctgcagaagtctctctggatgagggcccaaaagcaag 60
|||||
QY 93 aaaaattactaaagagaggtttgccaatgaatttctgaagctgaaaaggaatctaccacag 152
|||||
Db 61 aaaaattactaaagagaggtttgccaatgaatttctgaagctgaaaaggaatctaccacag 120
|||||
QY 153 tacaaggcgacacaaaacctctctacaactgtgctgagaatgccagaataatcaagaaa 212
|||||
Db 121 tacaaggcgacacaaaacctctctacaactgtgctgagaatgccagaataatcaagaaa 180
|||||
QY 213 aacagataaagatattttgcccctatgattatagccgggtgagaactatccctgataaac 272
|||||
Db 181 aacagataaagatattttgcccctatgattatagccgggtgagaactatccctgataaac 240
|||||
QY 273 tctgatgagattcccagctacatcaatgccaaacttcattaaaggagtttatggaaccaag 332
|||||
Db 241 tctgatgagattcccagctacatcaatgccaaacttcattaaaggagtttatggaaccaag 300
|||||
QY 333 gcttatattgccaccacagggtctcttatctacaacctctctggacttctctggagatgatt 392
|||||
Db 301 gcttatattgccaccacagggtctcttatctacaacctctctggacttctctggagatgatt 360
|||||
QY 393 tgggaataatagtgctcattatctgtatgcatgcatggatgataaaatgggaaagaaa 452
|||||
Db 361 tgggaataatagtgctcattatctgtatgcatgcatggatgataaaatgggaaagaaa 420
|||||
QY 453 aagtgtagcgtactggcctgagccagagagatgcagctggaatttgccctctctct 512
|||||
Db 421 aagtgtagcgtactggcctgagccagagagatgcagctggaatttgccctctctct 480
|||||
QY 513 gtatcctgtgaagctgaaaaaagaaaatctgtattataatcaggagactctaaagttaag 572
|||||
Db 481 gtatcctgtgaagctgaaaaaagaaaatctgtattataatcaggagactctaaagttaag 540
|||||
QY 573 ttcaatagtgaaactcgaactatctaccagtttcattacaagaatggccagaccatgat 632
|||||
Db 541 ttcaatagtgaaactcgaactatctaccagtttcattacaagaatggccagaccatgat 600
|||||
QY 633 gtaccttcattatagaccctattcttgagctcatctggatgtacgttgtttaccacagag 692
|||||
Db 601 gtaccttcattatagaccctattcttgagctcatctggatgtacgttgtttaccacagag 660
|||||
QY 693 gatgacagtggttcccatatgcattcactgctgctggctggtggaaggagactggtgtatt 752
|||||
Db 661 gatgacagtggttcccatatgcattcactgctgctggctggtggaaggagactggtgtatt 720
|||||
QY 753 tgtgctattgttgattatatacatggtatgtctaaaagatgggataatctctcgagaacttc 812
|||||
Db 721 tgtgcta---ttgattatatacatggtatgtctaaaagatgggataatctctcgagaacttc 777
|||||
QY 813 agtgttttcagtttgatccgggaaaatcggaacagagagccttcattagttcaaacacag 872
|||||
Db 778 agtgttttcagtttgatccgggaaaatcggaacagagagccttcattagttcaaacacag 837
|||||
QY 873 gaacaatatgaactggtctacaatgctgtatttagaactatttaagagacagatggtatggt 932
|||||
Db 838 gaacaatatgaactggtctacaatgctgtatttagaactatttaagagacagatggtatggt 897
|||||
QY 933 atcagagataaaacattctggaacagagatgcaagcaaacattgtattctctgagaaaaat 992
|||||
Db 898 atcagagataaaacattctggaacagagatgcaagcaaacattgtattctctgagaaaaat 957
|||||

QY 993 cacactctcaagcagactcttattctcttaattttaccacaaagtaccacaaagcagca 1052
DB 958 cacactctcaagcagactcttattctcttaattttaccacaaagtaccacaaagcagca 1017
QY 1053 aaaaatgatgaacacaaagcagacaaatggaatcaaaagaatcttctctcttgacttt 1112
DB 1018 aaaaatgatgaacacaaagcagacaaatggaatcaaaagaatcttctctcttgacttt 1077
QY 1113 aggaactctgaaataagtgcaaaaagagagctagttttgcacccctgctaaatcaagcact 1172
DB 1078 aggaactctgaaataagtgcaaaaagagagctagttttggcccccgg---aaaaccaggct 1134
QY 1173 tcttttgacttctgagctgaataatttgacaaaatctgacacaaaccatgaaa 1232
DB 1135 tcttttgacttctgagctgaataatttgacaaaatctgacacaaaccatgaaa 1192
QY 1233 tggcagacaaagcatttccaaatagttgggagcctcttcagaaagcatcaaaagtgtggat 1292
DB 1193 tggcag-caaaagccttccaaatag- -gggaagcctcttcagaaagcatcaaa- -tgaat 1248
QY 1293 tgggctctctttgtttgagggatgtcttaattctaaacctgttaaatgcagcaggaaga 1352
DB 1249 tgggctctctttgttt---aggaatgtcttaattctaaacctgt---atgcagcaggaaga 1303
QY 1353 tatttaattcaaaagtgccaataacacgacacaaatcaactccttttgaattgatcacag 1412
DB 1304 tattt---aatccaaagtgccaat-acccggaccacaaatcaact-cttttgaattgacccag 1359
QY 1413 cagagagaacaaagaggtggacagcaaggaacacttttcttatttggaaatcacaacca 1472
DB 1360 ca---agagaacaaaggtgacagcaagaaaa---ctttctatttgaatcacaacca 1413
QY 1473 catgattctgtttttagagatgcaggctcaaaaagtaaatgcattgttttcttcagcagaa 1532
DB 1414 catgattctgtttttagagatgcaggctcaaaaagtaaatgcattgttttcttcagcagaa 1473
QY 1533 ctgaattattcactgcatactactcaacacaaatacgttaatgcctctaagttaaaag 1592
DB 1474 ctgaattattcactgcatactactcaacacaaatacgttaatgcctctaagttaaaag 1533
QY 1593 caccatgactagtgctcttggtgtatattcttaccataccttttagtggaaaatccttat 1652
DB 1534 caccatgactagtgctcttggtgtatattcttaccataccttttagtggaaaatccttat 1593
QY 1653 ttttcatatgctcccaagtggtaccaggtcttaagatgtctcttgatttaccctgagaag 1712
DB 1594 ttttcatatgctcccaagtggtaccaggtcttaagatgtctcttgatttaccctgagaag 1653
QY 1713 caagatggaaactgttttctctctctctgttgccaacatcctctacatccctctctct 1772
DB 1654 caagatggaaactgttttctctctctctgttgccaacatcctctacatccctctctct 1713
QY 1773 tattacaattcacatagttctttatcactgaattctcccaacaaatatttctcactattg 1832
DB 1714 tattacaattcacatagttctttatcactgaattctcccaacaaatatttctcactattg 1773
QY 1833 aaccaggagtcagctgtactagcaactgtctcccaagatagatgataaaatccccctccca 1892
DB 1774 gaacaggagtcagctgtactagcaactgtctcccaagatagatgataaaatccccctccca 1833
QY 1893 ctctcctgacagacactgaatactatttattgtgttgaggaaagctgagaattctcacca 1952
DB 1834 tctcctgacagacactgaatactatttattgtgttgaggaaagctgagaattctcacca 1893
QY 1953 aatgttcccaaatcccttactcctcagctgtgaaaggtataaaattggaacatcactgggaatg 2012
DB 1894 aatgttcccaaatcccttactcctcagctgtgaaaggtataaaattggaacatcactgggaatg 1953
QY 2013 ggtggaaactctgaacacaaagaaattttagtactctgtgatacttagacacaaagcaagat 2072
DB 1954 ggtggaaactctgaacacaaagaaattttagtactctgtgatacttagacacaaagcaagat 2013

QY 2073 gtaaaactcgaagctcttaaatcagtgtaaaaaa 2105
DB 2014 gtaaaactcgaagctcttaaatcagtgtaaaaaa 2046

RESULT 5

AAF21801

ID AAF21801 standard; DNA; 1286 BP.

XX AAF21801;

XX 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 188.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
XX neurotropic; neurproliferative; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antinflammatory; antitumor; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
XX Addison's disease; allergy; autoimmune haemolytic anaemia;
XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
XX cardiovascular disorder; wound healing; neurological disease; ds.

OS Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI: 2000-611515/58.

XX P-PSDB; AAB58898.

XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention,
XX treatment and diagnosis of cancer, immune disorders, cardiovascular
XX disorders and neurological diseases

XX Claim 1; Page 622; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX invention. The breast and ovarian cancer associated DNA, protein, agonist
XX or antagonist sequences exhibit cytostatic; immunosuppressive;
XX neurotropic; neuroproliferative; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antinflammatory; antitumor; anticonvulsant;
XX antibacterial; antifungal; antiparasitic and cardiant activity. The
XX polynucleotide and protein sequences are used in the diagnosis of cancer,
XX particularly breast and ovarian cancer. The nucleic acid sequences,
XX proteins, agonists and antagonists may also be used in the diagnosis,
XX prevention and treatment of immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; cardiovascular disorders such as
XX myocardial ischaemias; wound healing; neurological diseases such as
XX cerebral anoxia and epilepsy; and infectious diseases.

SQ Sequence 1286 BP; 410 A; 273 C; 209 G; 389 T; 5 other;

Query Match 26.4%; Score 621.4; DB 21; Length 1286;
Best Local Similarity 98.7%; Pred. No. 4.2e-154;


```
XX DE Rat protein tyrosine phosphatase PRP20 cDNA.
XX KW Protein tyrosine phosphatase; PTP20; rat; signal transduction;
XX KW cell differentiation; cancer; neural injury; therapy; ss.
XX OS Rattus sp.
XX FH Key Location/Qualifiers
XX FT CDS 28..1389
XX FT polyA_signal 2150..2156
XX FT /*tag= a
XX FT /*tag= b
XX PN W09748723-A2.
XX PD 24-DEC-1997.
XX PF 17-JUN-1997; 97WO-IB00946.
XX PR 19-DEC-1996; 96US-0034286.
XX PR 17-JUN-1996; 96US-0019629.
XX PR 09-AUG-1996; 96US-0023485.
XX PR 13-NOV-1996; 96US-0030860.
XX PR 15-NOV-1996; 96US-0030964.
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PI Aoki N, Chen Z, Kharitonkov AI, Kim YW, Nayler O;
XX PI Ullrich A, Wang HY;
XX DR WPI; 1998-120302/11.
XX DR P-PSDB; AAW49906.
XX PT New phosphatase and kinase enzyme(s) - useful in the diagnosis and
XX PT treatment of signal transduction disorders
XX PS Claim 3; Fig 1a-b; 138pp; English.
XX CC This cDNA clone codes for a novel rat protein tyrosine phosphatase,
XX CC designated PTP20 (see AAW49906), that regulates growth factor
XX CC stimulation of cellular differentiation. The clone was isolated
XX CC from a rat pheochromocytoma PC12 cDNA library by PCR amplification
XX CC using primers based on consensus sequences (see AAW49915-16) of known
XX CC PTPs. The invention relates to novel proteins (see AAW49906-14)
XX CC involved in cellular signal transduction and to the nucleic acids
XX CC (see AAW17097-99) coding for them, and provides vectors, host cells,
XX CC purified recombinant proteins, methods for identifying compounds
XX CC that bind the novel proteins or abrogate their interactions with
XX CC natural binding partners, as well as methods for the diagnosis and
XX CC treatment of diseases associated with the novel proteins. For
XX CC PTP20, activators may act as anti-cancer therapeutics that
XX CC stimulate cell differentiation rather than proliferation, while
XX CC inhibitors may be useful for treating neural injuries by delaying
XX CC the differentiation of transplanted neuronal stem cells until they
XX CC are firmly grafted.
XX SQ Sequence 2226 BP; 574 A; 591 C; 604 G; 457 T; 0 other;

Query Match 9.1%; Score 214.2; DB 19; Length 2226;
Best Local Similarity 57.4%; Pred. No. 2.8e-46;
Matches 406; Conservative 0; Mismatches 298; Indels 3; Gaps 1;

QY 207 aagaaacacagatataagagatatatttgccctatgattatagccgggtagaactatccctg 266
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 aagaagaacgcgtacaaagcgtggaccgtagatgagacagagadtcactctccctg 258
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 ataacctctgtaggagatccctacatcaatgccaaacttcatttaaggaggttatgga 326
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 ctccaggagaagacacagagagattacattatgccaatttcacccggggcacagatgga 318
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 cccaaggttatattgccaccgccgggtcttttatctacaacctctcctgactctggagg 386
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 319 agccaggcctacattgcgacgcaaggaccctcctcactctgttgactctggcgc 378
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 387 atgatttgggaatatagtgctcttatctattgttatggcatgcatgagataaatggga 446
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 379 ctggtttgggagtttggaaatcaagtgatcttgatggcctgcaggadacagaataatgga 438
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 447 aagaaaaagtgagcgtactactggtgagccaggagagatgcagctggaatttggccct 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 439 cggaggaagtgtaacgctactggtgcccaggagcgggagcctctacaggccggcctttc 498
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 507 ttctcttatcctgtgaagctgaaaaaggaatctgattatataatcaggagctctaaaa 566
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 499 tgcataccctgacaaaaggagacagcactgactctcgacatcactctcaggacctccag 558
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 567 gtttaagtcaatagtgaaactcgaactatctaccagtttctaccagaatggccagac 626
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 559 gttatctccagaagaatcccgctctgtgccaccgctacagtagcatgtcttggcgggac 618
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 627 catgatgtacctctcatatagaccctattcttgagctcatctgggatgtacgtttgtac 686
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 619 caagggttcccagcagttccgatacacattctcaccatggtggagaggccgttgcttc 678
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 687 caagaggatgacagtggttcccatatgcattcactgcagtgctggtggaaggactgt 746
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 679 caaggacttgacctggaccctctgtgccactgcagtgctggtggtggaacaggt 738
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 747 gttatttgcattgttgattatcatatgattgttgctaaaagatgggataattcctgag 806
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 739 gctctgtgtgc---tggtgattacgtggagcagtgctctgactcagacaaatccccc 795
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 807 aacttcagtggtttcagtttcggtggaatgcggacacagagccttcattagttcaa 866
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 796 aatttcagcctcttgaagtgtctcgtggagatgcgaaacacgacacctgcagcgtgacg 855
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 867 acgcgggaacaatatgaactggtctacaaatgctgtattagaactatt 913
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 856 acagaggagcagtagcaggttctctgtaccacacagtggtcagctcatt 902
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
AAVI7099
ID AAVI7099 standard; cDNA; 2810 BP.
AC AAVI7099;
XX 20-JUL-1998 (first entry)
DT Human brain derived phosphatase 1 (BDP-1) cDNA.
XX Brain derived phosphatase 1; BDP-1; human; receptor;
XX protein tyrosine phosphatase; signal transduction;
XX therapy; diagnosis; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT misc_feature 31..43
XX FT /*tag= a
XX FT /*note= "GC-rich track, part of Kozak sequence"
XX FT CDS 44..1420
XX FT /*tag= b
XX FT /*transl_except= (pos:69..71, aa:Arg)
XX FT /*transl_except= (pos:162..164, aa:His)
XX FT /*transl_except= (pos:1182..1184, aa:Ser)
XX FT intron 733
XX FT /*tag= c
XX FT /*note= "site of 367 bp intron"
XX FT intron 799
XX FT /*tag= d
XX FT /*note= "site of 80 bp intron"
XX FT intron 878
XX FT /*tag= e
```

FT	misc_feature	/note= "site of 91 bp intron"
FT		2532..2545
FT		/*tag= f
FT		/note= "T-rich sequence required for
FT		polyadenylation"
FT		2740..2747
FT		/*tag= g
XX		
XX	polyA_signal	
XX		
PN	WO9748723-A2.	
XX		
PD	24-DEC-1997.	
XX		
XX	17-JUN-1997;	97WO-IB00946.
XX		
PR	19-DEC-1996;	96US-0034286.
PR	17-JUN-1996;	96US-0019629.
PR	09-AUG-1996;	96US-0023485.
PR	13-NOV-1996;	96US-0030860.
PR	15-NOV-1996;	96US-0030964.
XX		
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.	
XX		
PI	Aoki N, Chen Z, Kharitonov AI, Kim YW, Nayler O;	
PI	Ullrich A, Wang HY;	
XX		
DR	WPI; 1998-120302/11.	
DR	P-PSDB; AAW49908.	
XX		
PT	New phosphatase and kinase enzyme(s) - useful in the diagnosis and	
PT	treatment of signal transduction disorders	
XX		
PS	Claim 3; Fig 3a-d; 138pp; English.	
XX		
CC	This cDNA sequence codes for a novel human protein tyrosine	
CC	phosphatase (PTP), designated brain derived phosphatase 1 (BDP-1,	
CC	see AAW49908), that is expressed in most tissues and cell lines at	
CC	basal level, but is expressed high in epithelium origin cell lines	
CC	and cancer cell lines. BDP-1 was originally identified in a human	
CC	brain cDNA library, although the full-length clone was isolated	
CC	from the haematopoietic MEG01 cDNA library. The invention relates	
CC	to novel proteins (see AAW49906-14) involved in cellular signal	
CC	transduction and to the nucleic acids (see AAV7097-99) coding for	
CC	them, and provides vectors, host cells, purified recombinant	
CC	proteins, methods for identifying compounds that activate or	
CC	inhibit the novel proteins, as well as methods for the diagnosis	
CC	and treatment of diseases associated with the novel proteins.	
XX		
SQ	Sequence 2810 BP; 700 A; 808 C; 808 G; 494 T; 0 other;	
Query Match 8.6%; Score 202.2; DB 19; Length 2810;		
Best Local Similarity 56.5%; Pred. No. 4.7e-43;		
Matches 397; Conservative 0; Mismatches 303; Indels 3; Gaps		
QY	199	agaatatcaagaagaacagataaaggatattttgcctctgattatgacgggttagaac 258
Db		
QY	207	agaacgtgaggagaacccgctacaagacgtgctgcttatgatcagacgcgagtaacc 266
QY	259	tatccctgataaacctctgatgaggattccagctacatacaatgccaaacttcattaaggag 318
Db		
QY	319	tttatggaccacaaggcttatattgcccaccagggtcctttatctacaacccctctggact 378
Db		
QY	327	tggatggaacgtggctacattgccaagcgaaggacccttgctcacacccctgctagact 386
Db		
QY	379	tcctggaggatgatttgggaatatagtgctcttatcatattgtatggcatgcattggagatg 438
Db		
QY	387	tcctggagactggtctgggagtttggggtcaagggtgactctgatggcctgtcgcagagatag 446
Db		
QY	439	aaatcggaagaagaaagctgtagcgtacttgggctgagccagagagatgcagctggaat 498
Db		
QY	447	agaatggcgcgaagaaggttgagcgggtactggggccacagagcagcagcactgcagactg 506
Db		

XX	Sequence 4965 BP; 1683 A; 857 C; 971 G; 1454 T; 0 other;	
SQ		
	Query Match 5.5%; Score 129; DB 22; Length 4965; Best Local Similarity 53.2%; Pred. No. 1.4e-23; Matches 327; Conservative 0; Mismatches 275; Indels 13; Gaps 2;	
QY	164 caaaacctctacaaactgtgctgagaaatccaagaatacaagaaaaacagatataa 223	XX
Db	1729 cagcaagttccctataaagaagccttaaccagaataaaaaccgttatgt 1788	XX
QY	224 ggatatttgcctatgattatagccgggtagaaactcctcgtataaccctctgatgagga 283	XX
Db	1789 tgacattctcctatgattataaccgtgtgaactctgagataaacggagatcgagg 1848	XX
QY	284 ttccagctacataaagccaacttcattaaaggagtttatggacccaagcttatattgc 343	XX
Db	1849 gtcaaaactacataaagccagctattatggtttcaagaacccaggaaaatacattgc 1908	XX
QY	344 caccaggggtccctttatcacaaacctcctggactctgagagatgatttggaaatatag 403	XX
Db	1909 tgcaaaaggtcccaaggatgaactgttgatgtattctcgtggaggtatttgggaacagaa 1968	XX
QY	404 tgccttatacttattgttgcgcgtgagtgatgaatgggaaagaaaagtgtgagcg 463	XX
Db	1969 agccacagttattgtcatgtcactcgatgtgaagaagaaacagaaacaagtgtgcaga 2028	XX
QY	464 ctactgggctgagccaggagagatgcagctggaatttggcccttctctgtatcctctgtga 523	XX
Db	2029 atactggcgtcaatggaagaggcactcggcgttttggagagtggttgttaaaatct 2088	XX
QY	524 agctgaaaaaggaatact----gattataatcaggactctaaaagttaagtccaata 579	XX
Db	2089 aaccagcaaaagatgtcctagatatacattcagaaattgaacattgtaatacaaa 2148	XX
QY	580 gtgaactgaaact-----atcacagtttctattacaagaattggccagaccatg 630	XX
Db	2149 aagaaaaagcaactggaagagggtgactcacattcagttccaccagtgccagaccag 2208	XX
QY	631 atgtactcctcaatagaccctattcttgagctcactcgggatgtacgttgttaccag 690	XX
Db	2209 ggtgctcaggatcctcactgtcctcctcaaaactgagaagagagtgaaatgccttcagca 2268	XX
QY	691 aggatgacagtggtcccatatgcattcactgcagtgctgctgtggaagactggtgtta 750	XX
Db	2269 atttctcagtggtcccaattggtgctgagtgctggtgtgtgtggcgacacaggaaact 2328	XX
QY	751 tttgtgctattgttg 765	XX
Db	2329 atatcggaattgatg 2343	XX
	RESULT 14	
AA160204	AA160204 standard; cDNA; 4968 BP.	
AC	AA160204;	
AC	AA160204;	
DT	22-OCT-2001 (first entry)	
XX	Human polynucleotide SEQ ID NO 4193.	
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	

XX	26-JUL-2001.	XX
XX	26-DEC-2000; 2000WO-US34263.	XX
XX	21-JAN-2000; 2000US-0488725.	XX
PR	25-APR-2000; 2000US-0552317.	XX
PR	09-JUL-2000; 2000US-0598042.	XX
PR	19-JUL-2000; 2000US-0620312.	XX
PR	03-AUG-2000; 2000US-0653450.	XX
PR	14-SEP-2000; 2000US-0662191.	XX
PR	19-OCT-2000; 2000US-0693036.	XX
PR	29-NOV-2000; 2000US-0727344.	XX
XX	(HYSE-) HYSEQ INC.	XX
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	XX
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	XX
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	XX
XX	WPI: 2001-442253/47.	XX
DR	P-PSDB; AAM41048.	XX
XX	Novel nucleic acids and polypeptides, useful for treating disorders	XX
PT	such as central nervous system injuries -	XX
XX	Claim 1; SEQ ID NO 4193; 10078pp; English.	XX
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and	XX
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,	XX
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	XX
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	XX
CC	of the invention may be used to treat diseases of the peripheral nervous	XX
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	XX
CC	localised neuropathies and central nervous system diseases, such as	XX
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	XX
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	XX
CC	utilisation of the activities such as: Immune system suppression,	XX
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	XX
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	XX
CC	assays for receptor activity, arthritis and inflammation, leukaemia and	XX
CC	C.N.S disorders.	XX
CC	Note: The sequence data for this patent did not form part of the printed	XX
CC	specification.	XX
XX	Sequence 4968 BP; 1678 A; 861 C; 974 G; 1455 T; 0 other;	XX

QY	164 caaaacctctacaaactgtgctgagaatgccaaagaataatcaagaaaaacagatataa 223	
Db	1730 cagcaagttcctataaaggaaagctgcgaagccctttaaaccagaataaaaaccgttatgt 1789	
QY	224 ggatatttgcctatgattatagccgggtagaaactatccctgataaacctcgtatgagga 283	
Db	1790 tgacattctcctatgattataaccgtgtgaactctgagataaacggagatcgagg 1849	
QY	284 ttccagctacataaagccaacttcattaaaggaggttatggacccaagcttatattgc 343	
Db	1850 gtcaaaactacataaagccagctatatattggtttcaagaacccaggaaaatacattgc 1909	
QY	344 caccaggggtcctttatctacaaacctcctcgtgactcttgaggatgatttgggaatatag 403	
Db	1910 tgcaaaaggtcccaaggatgaactgttattcttgaggatgatttgggaacagaa 1969	
QY	404 tgccttatacttattgttgcgcgtatgagtgatgaatgggaaagaaaagtgtgagcg 463	
Db	1970 agccacagttattgtcactcgtatgtgaagaagaaacaggaacagtggtgcaga 2029	
QY	464 ctactgggctgagccaggagagatgcagctgggaatttggcccttctctgtatcctctgtga 523	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 10:35:25 ; Search time 54.75 seconds
(without alignments)

308.722 Million cell updates/sec

Title: US-09-600-358A-4

Perfect score: 3615

Sequence: 1 MPQREILQFLDAQSKKIT.....RPSKVKLRSPKSGKNFSL 692

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3543.5	98.0	807	4	US-09-081-345-2
2	2356.5	65.2	802	4	US-09-081-345-18
3	1345.5	37.2	278	4	US-08-821-278A-18
4	1008.5	27.9	272	4	US-08-821-278A-19
5	928.5	25.7	223	2	US-08-885-992-20
6	928.5	25.7	223	2	US-09-144-925-20
7	748	20.7	453	3	US-08-951-260A-7
8	736.5	20.4	453	4	US-08-821-278A-2
9	508	14.1	1337	3	US-08-854-585-2
10	508	14.1	1337	5	PCT-US95-05512-2
11	492.5	13.6	277	2	US-08-885-992-22
12	492.5	13.6	277	2	US-09-144-925-22
13	492.5	13.6	1439	2	US-08-449-644-2
14	492.5	13.6	1439	2	US-08-087-244A-2
15	487	13.5	595	1	US-08-202-389-6
16	485.5	13.4	593	1	US-08-018-129-5
17	485.5	13.4	593	2	US-08-448-250-5
18	485	13.4	1442	1	US-08-015-986A-3
19	485	13.4	1442	2	US-08-446-363-3
20	484	13.4	631	1	US-08-202-389-8
21	479	13.3	1452	2	US-08-652-971-4
22	479	13.3	1452	2	US-08-991-258A-4
23	479	13.3	1452	2	US-08-769-399-4
24	479	13.3	1452	3	US-08-991-953A-4
25	478.5	13.2	1457	2	US-08-652-971-3
26	478.5	13.2	1457	2	US-08-449-644-1
27	478.5	13.2	1457	2	US-08-087-244A-1

Query Match 98.0% ; Score 3543.5 ; DB 4 ; Length 807 ;

Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 8, Appli
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-081-345-2
; Sequence 2, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; APPLICANT: Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ For Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081.345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-345-2

```
Best Local Similarity 99.4%; Pred. No. 3e-291;
Matches 681; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDQREILQFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKNRY 60
Db 1 MDQREILQFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAEKPKNIKNRY 60

QY 61 KDILPYDYSRVLSLTSDEDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDYSRVLSLTSDEDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120

QY 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180

QY 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180

QY 181 ETRTIYQFYHKNPDPHDVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
Db 181 ETRTIYQFYHKNPDPHDVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240

QY 241 VDTWMLLDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVAVLELFRQMDVIRD 300
Db 241 -DVTWMLLDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVAVLELFRQMDVIRD 299

QY 301 KHSGETSOAKHCIPKHNHLOADSYPNLPKSTTKAAKMNQOQRTKWEIKESSDFPRTS 360
Db 300 KHSGETSOAKHCIPKHNHLOADSYPNLPKSTTKAAKMNQOQRTKWEIKESSDFPRTS 359

QY 361 EISAKELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQSIDLGS 420
Db 360 EISAKELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQSIDLGS 419

QY 421 LLFEGCSNKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDSKENFSYLESQPHDS 480
Db 420 LLFEGCSNKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDSKENFSYLESQPHDS 479

QY 481 CFVEMQAQKVMHVSSALNYSLPYDSKHQIRNASNVKHHDSALGYVSYIPLVENVYFSS 540
Db 480 CFVEMQAQKVMHVSSALNYSLPYDSKHQIRNASNVKHHDSALGYVSYIPLVENVYFSS 539

QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLNSLPTNISLLNQE 600
Db 540 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLNSLPTNISLLNQE 599

QY 601 SAVLATAPRIDDEIPLPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWEGGT 660
Db 600 SAVLATAPRIDDEIPLPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGTSLWEGGT 659

QY 661 SEPKKFDDSVILRPSKSVKLRSKPS 685
Db 660 SEPKKFDDSVILRPSKSVKLRSKPS 684

RESULT 2
US-09-081-345-18
; Sequence 18, Application US/09081345
; Patent No. 6228641
; GENERAL INFORMATION:
; APPLICANT: Bahija Jallal
; APPLICANT: Gregory D. Plowman
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: PTP04 RELATED DISORDERS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; SUITE: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/081,345
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/047,222
; FILING DATE: May 20, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 234/253
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 802 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-081-345-18

Query Match 65.2%; Score 2356.5; DB 4; Length 802;
Best Local Similarity 68.5%; Pred. No. 8.5e-191;
Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

QY 1 MDQREILQFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAENAKNIKNRY 60
Db 1 MDQREILQFLDQAQSKKTKKEEFANFLKLRQSTKYKADKTYPTTVAEKPKNIKNRY 60

QY 61 KDILPYDYSRVLSLTSDEDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120
Db 61 KDILPYDYSRVLSLTSDEDSYINANFIKGVYGPAYATQGPLSTLLDFWRMIWEY 120

QY 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180
Db 121 SVLIIVMACMEYEMGKKKERYWAEFGEMOLEPFPSPVSCAEAKRKSXYIIRTLKVKFNS 180

QY 181 ETRTIYQFYHKNPDPHDVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 240
Db 181 ETRTIYQFYHKNPDPHDVPSSIDPILLELWVRCYQEDDSVPICIHCSAGCGRTGVICAI 239

QY 241 VDTWMLLDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVAVLELFRQMDVIRD 300
Db 240 VDTWMLLDGIIIPENFVSFLIREMRTQPSLVQTOEQYELVYNVAVLELFRQMDVIRD 299

QY 301 KHSGETSOAKHCIPKHNHLOADSYPNLPKSTTKAAKMNQOQRTKWEIKESSDFDF 357
Db 300 NHLGRIQACSIPEQSLVEADSCPLDLPKNAMRDVKTTHQHSKQGAESGSSGL 359

QY 358 RTSEISAKEELVLHPAKSSTSFDFLELNTSFDKNADTTMKWTKAPPIVGEPLQKHQS 417
Db 360 RTSTMNAEEELVLSAKSSPFCNLCNGCNKNAVITRNGQARASPVVGEPLQKYQSLD 419

QY 418 LGSLLFEGCSNKPVNAAGRYFNSKVPITRTKSTPFELIQORETKEVDSKENFSYLESQ 477
Db 420 FGSMLFGCSPSALPINTADRYHNSKGPVKRTKSTPFELIQORETNDLAVGDGFSQ 479

QY 478 HDSCFV-EMQAQKVMHVSSALNYSLPYDSKHQIRNASNVKHHDSALGYVSYIPLV 536
Db 480 HEHYSRELVQVQVAHVSSLEELNYSLP-----CACDASCVPVPHSPGALRVHLYTSLAEDP 534

QY 537 YFSSWPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNSHSSLNSLPTNIS 596
Db 535 YFSSSPNSADSMSFDPLEKQDGTSPGALLPASSTTFFYSNPHDLSVMTLTSFSP 594

QY 597 LNOESAVLATAPRIDDEIPLPLPVRTPESTIVVEEAGEFSPNVPKSLSSAVKVKIGT 656
```

```

Db      595   LNOETAVEAPSRRDTDEIPPLPERTPEFIVVEEAGEPSRVTESL--PLVVTGASPE 652
          |||||:: : ||||||| ||||||| ||||| : || : | | |
QY      657   WGTSEPKFKDDSVILRPKSKVKLRSPKSGKN 688
          ||||| ||||| ||||| ||||| ||||| ::
Db      653   CSCTSBMKS-HDSVGFTPKNVKLRSKSDRH 683
          ||||| ||||| ||||| ||||| ||||| ::

```

```

RESULT      3
US-08-821-278A-18
; Sequence 18, Application US/08821278A
; Patent NO. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 18
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo Sapien
US-08-821-278A-18

```

Query Match	37.2%;	Score 1345.5;	DB 4;	Length 278;
Best Local Similarity	88.5%;	Pred. No. 5.6e-106;		
Matches	247;	Conservative 18;	Mismatches 13;	Indels 1; Gaps 1;
QY	24	FANEFLLKRSQSYKADKYPTTVAENAKNIKNRYKDILPYDYSRVLSLITSDESS	83	
Db	1	FASEFLKRSQSYKADKYPTTVAQRPNIKKNRYKDILPYDHSVLSSLTSDESS	60	
QY	84	YINANFLKGYVGPKEYATATQGPLSTTLTLLDFRMITWEYSLVLIIVMACHEYEMGKKCCRYW	143	
Db	61	YINASF.LKGYVGPKEYATATQGPLSTTLTLLDFRMITWEYSLVLIIVMACFEFEMGKKCCRYW	120	
QY	144	AEPCEMOLERGPFSVSCAEAKKRSDYIIRTLKVKFNSETRTIYOFHYKKNPDHVPSSID	203	
Db	121	AEPGETOLQGPFSISCAEAKKSDYKIRTLKAKFNNETRIIYOFHYKKNPDHVPSSID	180	
QY	204	PILELWDRVCYQEDDSVPICIHCSACGRTGVCATVDYTMLLKDGIIIPENFSVFLI	263	
Db	181	PIQLWDMRCYQEDDCVPICIHCSACGRTGVICA-VDYTMLLKDGIIIPKNEVFENLI	239	
QY	264	REMRTORPSLVQTEQYVELVYNAVLELFKKRQMDVIRDKH	302	
Db	240	OEMRTORPSLVQTEQYVELVYSAVLEFKRHMVDISNH	278	

```

RESULT      4
US-08-821-278A-19
; Sequence 19, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 19
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo Sapien
US-08-821-278A-19

```

Query Match	27.9%	Score 1008.5;	DB 4;	Length 272;
Best Local Similarity	65.8%	Pred. No. 1.8e-77;		

Matches	179;	Conservative	46;	Mismatches	46;	Indels	1;	Gaps	1;
QY	24	FANEFLKLRQSTKYRADTKTYPTTVAENAKNIKKNRYKDILPYDSRVLSLTSDSDS	83						
		I I : : : I I I I I : : I I I : : I I I I I : : I I I I I : : I I I							
Db	1	FARDFWRLRLSPKYRTEKIYPTTGKEKENVKNNRYKDILPDHRSVRLTLTKTPSQDSD	60						
QY	84	YINANFTKGVGPKAYIATOGPLSTLLDPWRMIWEYSVLIIYVMACMEYMGGKKCKERYW	143						
		I I							
Db	61	YINANFTKGVGPKAYIATOGPLANIVIDFWRWVETNVIIVYMAREFEMGRRCKERYW	120						
QY	144	AEPGMQLFEGPFSVSCEAEKRKSDYIIRTLKLVFNSETRTIYOFHYKNMPDHDPSSID	203						
		I I : : : I I I I I : : I I I I I : : I I I I I : : I I I I I : : I I I I I							
Db	121	PLYGEDPIFAFPKISCEDBQAQTDYFIRTLLEFONESRLYQFYHVNMPPDHDPSSFD	180						
QY	204	PILLELWDVRCYQEDSDVPICHCISAGCGRTGYCAIVDTMLLKDGIIIPENFSVFSLI	263						
		I I							
Db	181	SILDUMLSMKRYGEHEVDPICHCSAGCGRTGAICAI-DITWNLLKAGKIPEEFNFVNLI	239						
QY	264	REMTORPSLVQTQEYQELVYNNAVLELFKROM	295						
Db	240	OEMRTORHSAVOTKEOVYLVRHAIAOLFKEOL	271						
		I I							

```

RESULT      5
US-08-685-992-20
; Sequence 20, Application US/08685992
; Patent No. 5912138
;
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992
; FILING DATE: 25-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

Query Match 25.7%; Score 928.5; DB 2; Length 253;
Best Local Similarity 65.7%; Pred. No. 9.4e-71;
Matches 167; Conservative 41; Mismatches 43; Indels 3; Gaps 2;

Best Local Similarity 43.1%; Pred. No. 4.5e-55;
Matches 154; Conservative 58; Mismatches 117; Indels 28; Gaps 7;

QY 3 QREILQFLDEAQSCKITKEE-FANFELKLRKSTYKADKTYPTTVAENAKNKNRYK 61
Db 4 OSOLVRSFLEQQAARDHRKGAIIAREFSDIKARSAVKTEGVCSTKAGSOOGSKNRYK 63
QY 62 DILPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTLLDFWMIWEYS 121
Db 64 DVVPYDETRVILSLQEEGHGDIYINANFIRGTDSQAYIATQGPLHTLLDFWRLWFEFG 123
QY 122 VLIIVWACHMEYEMGKKCERYWAEPEGMQLEFGPFPFVSCEAEKR-KSDYIIRLKVKNFS 180
Db 124 IKVILMACQETENGRKRCERYWAQERE-PLQAGPFCITLTKETALTSIDILRTLQVTFQK 182
QY 181 ERTTIOFHYKNPDDHVPSSIDPILFELIWDVRCYQEDDSDVPICHSCAGCGRTGVICAI 240
Db 183 ESRPVHQLQYMSWPDHGVPSSSHILTMVEEARCLQGLGPGPLCVHCSAGCGRTGVICA - 241
QY 241 VDYTWMLKDGIIIPENFVSFLIREMRTQPSLSVQTOEQYELVYNNAVLELFKQMD---- 296
Db 242 VDYVROLTLTQTPNFSLFEVLEMRKQRPAAVQEEQYRFLYHTVAQLFSRTLQNNSP 301
QY 297 -----VTRDKHSCTESQAKHICEKNHTLQADSYS-----PNLPKSTTKA 336
Db 302 LYQNLAKENRAPICKDSSSLTSSA---LPATSRPLGGVLRSSISVPGPTLPMDATYA 355

RESULT 8
US-08-821-278A-2
; Sequence 2, Application US/08821278A
; Patent No. 6238902
; GENERAL INFORMATION:
; APPLICANT: Cheng, Jill
; TITLE OF INVENTION: Protein Tyrosine Phosphatases
; FILE REFERENCE: P1010R1
; CURRENT APPLICATION NUMBER: US/08/821,278A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 2
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Mus Musculus
US-08-821-278A-2

Query Match 20.4%; Score 736.5; DB 4; Length 453;
Best Local Similarity 44.0%; Pred. No. 4.2e-54;
Matches 144; Conservative 65; Mismatches 109; Indels 9; Gaps 6;

QY 5 EILQKFLDEAQSCKITKEE-FANFELKLRKSTYKADKTYPTTVAENAKNKNRYKDI 63
Db 6 DLVRSFLEQQAARDHRKGAIIAREFSDIKARSAVKTEGVCSTKAGSRLGNTNKNRYKDV 65
QY 64 LPYDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTLLDFWMIWEYSVL 123
Db 66 VADETRVILSLQEEGHGDIYINANFIRGTDSQAYIATQGPLHTLLDFWRLWFEVGK 125
QY 124 IIVWACHMEYEMGKKCERYWAEPEGMQLEFGPFPFVSCEAEKR-KSDYIIRLKVKNSET 182
Db 126 VILMACQETENGRKRCERYWAEQE-PLKAGPFCITLTKETLNADITLRTLQVTFQKEF 184
QY 183 RTYQFHYKNPDDHVPSSIDPILFELIWDVRCYQEDDSDVPICHSCAGCGRTGVICAI 242
Db 185 RSVHQLQYMSWPDHGVPSSSHILTMVEEARCLQGLGPGPLCVHCSAGCGRTGVICA -VD 243
QY 243 YTWMLKDGIIIPENFVSFLIREMRTQPSLSVQTOEQYELVYNNAVLELFKQMDVIRDKH 302
Db 244 YVROLTLTQTPNFSLFEVLEMRKQRPAAVQEEQYRFLYHTVAQLFSR---TLQDTS 300
QY 303 SGTESQAKHCIP--EKNHTLQADSYS 327
; Sequence 2, Application PC/TUS9505512

Db 301 PHYQNLKENCAPICKEAFSLRTSSALP 327

RESULT 9
US-08-854-585-2
; Sequence 2, Application US/08854585
; Patent No. 6114140
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,585
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,940
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-854-585-2

Query Match 14.1%; Score 508; DB 3; Length 1337;
Best Local Similarity 38.3%; Pred. No. 5.1e-34;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;

QY 15 QSKKITKEEFANFELKLRKST-----KYKADK----TYPTTVAENAKNKNRYKDILP 65
Db 1017 KSLIRVENFATFKKQOADSNCGFAEYEDLKLVGISQPKYAAELAKNKNRYNNVLP 1076
QY 66 YDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGPLSTLLDFWMIWEYSVLII 125
Db 1077 YDISRVKLSVQTHSTD-DYINANTMPCYHSHKDFIATQGPLNTLKDQFWMVWEKNVYAI 1135
QY 126 VMACMEYEMGKKCERYWAEPEGMQLEFGPFPFVSCEAEKRKSDYIIRLKVK--FNSETR 183
Db 1136 IMLTCKVEOGRTKCEEYW--PSKQAQDYGDITVAMTSEIVLPEWITRDTFTVKNIQTSESH 1193
QY 184 TIYQFHYKNPDDHVPSSIDPILFELIWDVRCY--QEDDSVPICHSCAGCGRTGVICAI 241
Db 1194 PLQGFHTSWPDHGVDPDTLLINFRYLVRDYMKQSPSPFVILVHCAGVGRGTFFAID 1253
QY 242 DYTWMMLKDGIIIPENFVSFLIREMRTQPSLSVQTOEQYELVYNNAVLELFKQMD 296
Db 1254 RLIVQIENENTV---DVYGVIVDLRMHRLPMVQTEQYVFLNCVLDIVRSQKD 1304

RESULT 10
PCT-US95-05512-2
; Sequence 2, Application PC/TUS9505512

GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K. and stman, Arne
; TITLE OF INVENTION: Density Enhanced Protein Tyrosine
; TITLE OF INVENTION: Phosphatase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05512
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/31954
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1337 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-05512-2

Query Match 14.1%; Score 508; DB 5; Length 1337;
Best Local Similarity 38.3%; Pred. No. 5.le-34;
Matches 113; Conservative 51; Mismatches 111; Indels 20; Gaps 7;

QY 15 QSKKITKEEFANFLKLRQST-----KYKADK---TYPTVAENAKNKKRYKDILP 65
DB 1017 KSKLRIVENFEAFYKQKQADSCNGFAEYEDLKVISOQPKYAAELAEKRGKRYNNVLP 1076
QY 66 YDYSRVLSLITSDSSYINANFIKGYGPKAYIATQGLSTLLDFWRMIWEYSVLII 125
DB 1077 YDISRVKLSVQTHSTD-DYINANYMPGYHSKKDFIATQGLPNTLKFWRWVWKNVYAI 1135
QY 126 VMACMEYEMGKKCYRWAEPMOLEFGPFPVSCEAEKRSYIIRTLKVK--FNSETR 183
DB 1136 IMLTRKCEQGRTKCEYW--PSKQADYGDITVAMTSEIVLPEWTRDFTVKNIQTSEH 1193
QY 184 TIYQHYKNKPDHPVSSITDPILELIWDVRCY--QEDDSVPICIHCSAGCGTGVCAIV 241
DB 1194 PLRQHFHTSWPHGVPDITDLINFRYLVDYMKOSPPSPILVHCSAGVGTGTFFIAID 1253
QY 242 DYTMLLKDGIIPEFVSFLIREMRTPQSLVQEQYELVYNVAVLELFKRQMD 296
DB 1254 RLIIQIENENTV---DVGIVYDLRMHRLMVQIEDQYVFLNQCVDIVRSQKD 1304

RESULT 11
US-08-685-992-22
; Sequence 22, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992
; FILING DATE: 25-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-685-992-22

Query Match 13.6%; Score 492.5; DB 2; Length 277;
Best Local Similarity 38.2%; Pred. No. 8.6e-34;
Matches 108; Conservative 40; Mismatches 92; Indels 43; Gaps 8;

QY 45 PTTVAENAK--NIKKNRYKDILPYDSRVSELSLITSDSSYINANFIKGYGPKAYIAT 102
DB 1 PSETSEGDKHHTSKNRYTNILPVNHNTRVOLKIQDKGSDYINANYIDGAY-PROFICT 59
QY 103 QGPLSTLLDFWRMIWEYSVLIIVMACMEYEMGKK---KCEYWAEP-----GEMQ 150
DB 60 QGPLNTIADFWRWVWENRCRIIIVMLSRSESESENCRIKCDRYWPEQIGGEQFSIYNGN 119
QY 151 LEFGPFPVSCEAEKRSYIIRTLKVKFNSETRTIYQHYKNKPDHPDVSSIDPILELIW 210
DB 120 EVFGTYSVELVEVIOCREIITRNIRLTFBGETRDTIQYQYEGWPDHNPDPHQPFRQLLH 179
QY 211 DVRCYQ-----EDDSVPICIHCSAGCGTGVCAIV-----DYTWMLLKDG 251
DB 180 SITRNQNIIPSSDRNVPILVHCSAGVGTGTCTAVIMMKKLDHYFKOLDYNSRI---- 235
QY 252 IIPENFVSFLIREMRTPQSLVQEQYELVYNVAVL-ELFKR 293
DB 236 ----DFNLFISIVLKLREQRPGMVQQLQYLFYCYKILDEIYHR 274

RESULT 12
US-09-144-925-22
; Sequence 22, Application US/09144925
; Patent No. 5951979
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.

STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02421-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,925
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,992
FILING DATE: July 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL96-03Z
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
TELEX:
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-144-925-22

Query Match 13.6%; Score 492.5; DB 2; Length 277;
Best Local Similarity 38.2%; Pred. No. 8.6e-34;
Matches 108; Conservative 40; Mismatches 92; Indels 43; Gaps 8;

Qy 45 PTTVAENAK--NIKKNRYKDLIPDYNSVELSLTSDSSYINANFIKGVYGPAYAT 102
Db 1 PSETSEGGKKNTSKNRYINLVNHTRVOLKTKODKSGSYINANYIDGAY-PKQFICT 59
Qy 103 QGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKK---KCRYWAEP-----GEMQ 150
Db 60 QGPLNTIADFWAWWENCRILIIWLSRESEGCNCRKCDRYWPEQIGGQFSIYGNGN 119
Qy 151 LEFGFVSVCBAERKSDYIIITLKVFNSETRTIYQPHYKNWPDHVPSSIDPILLELIW 210
Db 120 EVFGTYSVELVEVIOCREIITRNIRLTPEGETRDTIQYEGWPDHNPDPHTQPPQLLH 179
Qy 211 DVRCYQ-----EDDSVPICHCAGCGRTGVCIAV-----DYTWMLLKDG 251
Db 180 SITRQRIQIPSDRNVPPIVHCAGVGRGTCTFTAVIMMKKLDHYFKQLDYSNRI---- 235
Qy 252 IIPENFSVFLIREMRTORPSLVOTQEQYELVYNVAVL-ELFKR 293
Db 236 ----DFNLFSLVLRQRPWVQOOLEQYLFICYTKILDEIYHR 274

RESULT 13
US-08-449-644-2
Sequence 2, Application US/08449644
Patent No. 5856162
GENERAL INFORMATION:
APPLICANT: Schlusser, Joseph
APPLICANT: Sap, Jan M.
APPLICANT: Ulrich, Axel
APPLICANT: Vogel, Wolfgang
APPLICANT: Fuchs, Miriam
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-KAPPA
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,644
FILING DATE: 24-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/087,244
FILING DATE: 01-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-042
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1439 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-449-644-2

Query Match 13.6%; Score 492.5; DB 2; Length 1439;
Best Local Similarity 29.1%; Pred. No. 1.2e-32;
Matches 136; Conservative 77; Mismatches 167; Indels 87; Gaps 16;

Qy 21 KEFANEFLKLRSTYKADKTYPTVAENAKNIKKNRYKDLIPDYNSVELSLTSD 80
Db 888 KEYESF---EGOSASW-----DVAKKDONRAKNRYGNIITAYDHSRVILQVEDDP 936
Qy 81 DSSYINANFIKGVYGPAYIATQGPLSTLLDFWRMIWEYSVLIIIVMACMEYEMGKKCE 140
Db 937 SSDYINANYIDGQRPISHYIATQGPVHETVYDFWRMIWQEOSACIVMTNLVVEGRVKCY 996
Qy 141 RYWAEPGEMQLEFGFPFVSCEAEKRSYIIR--TLKYFNSETRTIYQPHYKNWPDHVD 198
Db 997 KYWPDDEV---YGDFTVCVEMEPLEAVVVRTFLERRGYNEIREVKQFHFTGWPDHGV 1053
Qy 199 PSSIDPILLELIWRCVOEDDSVPICHCAGCGRTGVCIAVDTWMLLKDGIIPEFNS 258
Db 1054 PYHATGLSLFRRVKLSNPPSAGPIVHCSAGAGRTGCVIIVIMLDMAREGVV-----D 1109
Qy 259 VFLIREMRTORPSLVOTQEQYELVYNVAVL-----FK-RQMDVIR----- 299
Db 1110 IYNCVKALRSRRINNVQTEQYIFIHDAILEACLCGETAIPVCEFKAAAYFDMIRIDSOTN 1169
Qy 300 DKHSGTESQAKH-----CIPEKNHTLIQ--ADSYSPN--LPKSTTTAAKMMN 341
Db 1170 SSKLKDEQTLNSVTPRLQAEDCSTACLP-RNHDKNRNMDMLPPDRCLPFLITIDGESSN 1228
Qy 342 QORTKM--EIKESSDFRTSEISAKEELVLHPAKSSTSFDFLELNYSF----- 388
Db 1229 YINAALMDSYRQPAAF-----IVTQYPLNTVKDFWRLVYDYGCTSIYMLNEV 1276
Qy 389 DKNADTTMKWOTKAF----PIVGEPLQKHQSLDGLSLLFEGCSNSKP 431
Db 1277 DLSQGCPOYWPEEGLMRYGPIQVECMSCMDCDVINFRICNLTTRP 1323

```

Qy 342 QORTKM--ETKESSSPDFTSEISAKEELVHLHPAKSSTSFDFLELNYSF----- 388
Db 1229 YINAALMDSYRQPAAF-----IVTQYPLPNTVKDFWRLVYDYGCTSIYMLNEV 1276
Qy 389 DKNADTTMKWQTAFAF-----PIVGEPLQKHQSLDGLSLLFEGCSNSKP 431
Db 1277 DLSGCGQYWPBEGMLRYGPIQVECMSCSMDCCDVINRIFRNCNLTRP 1323

RESULT 15
US-08-202-389-6
; Sequence 6, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: Freeman Jr., Robert M.
; APPLICANT: Plutzky, Jorge
; APPLICANT: Neel, Benjamin G.
; APPLICANT: Rosenberg, Robert D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,141
; FILING DATE: 31-JAN-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: BIH92-05MA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-202-389-6

```

	Query Match	13.5%	Score 487;	DB 1;	Length 595;
	Best Local Similarity	33.0%;	Pred. No. 8.4e-33;		
	Matches 128; Conservative	69;	Mismatches 129;	Indels 62;	Gaps 14;
Qy	15 QSKKITKEEFANEPLKLRQSTKYKADTYPTTVAENAKNNKNRYKDILPVDYSRVLS	74	:	:	:
	: : :	: : :	:	:	:
	: : :	: : :	:	:	:
Dd	235 ESDTKAGFWEFEESLQEQVK-----NLHRLGQRPNKGKKNYLLFPDHSRVLQ	290	:	:	:
	:	:	:	:	:
Qy	75 LITSD-EDSSYINANFIKG-VYGP----KAYIATQGPLSTLLDPWRMIWYSVLIIVMA	128	:	:	:
	:	:	:	:	:

Db 291 GRDSNIPGSDYINANYIKNQLLGPDENAKTYIASQGCLEATVNDFWQAMQOENSRIYMT 350
Qy 129 CMYEMGKKKCYWAEPEGMQLEFGPFSVSCAEKRKSDYIIRTLKVK--FNSE-TRTI 185
Db 351 TREVEKGRNKCVPYWEVG-MORAYGPYSVTNVGEHDTTEYKRLTLQVSPDLNDGLIREI 409
Qy 186 YOFHYKNMPDHDVPSSIDPILELIWDVRCYOED--DSVPICIHCSAGCGRTGVICAIVDY 243
Db 410 WHYQYLSWPDHGVSEPGVLSFLDQINORQESLPHAGPIIVHCSAGIGRTGTIIVI--- 466
Qy 244 TWMLLKDGIIIPENSVFSL-----IREMRTORPSLVOTQOYELVYNV---LELF 291
Db 467 -----DMLMENISTRGLDCDIDIQKTQIMVRAQRSGMVQTEAQYKFIYVAIAQFIETT 519
Qy 292 KROMDVIRDKHSGTESQAKHCIPEKNHTLOADSYSNLPKSTTKAAK-----M 339
Db 520 KKKLEVIQSQ-KGQSEYGN-----TTPPAKNNAHAKASRTSSKHKEDVYENL 567
Qy 340 MNQORTKMEIKESSSDFRTSEISAKEE 367
Db 568 HTKNKREEKVKQRSADKESKSGSLKRK 595

Search completed: May 26, 2002, 12:17:53
Job time: 6148 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2002, 10:34:06 ; Search time 119.16 Seconds
(without alignments)
645.041 Million cell updates/sec

Title: US-09-600-358a-4
Perfect score: 3615
Sequence: 1 MDQREILQFLDEAQSCKIT.....RPSKSVKLKSPKSGKNFSWL 692

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3615	100.0	692	20	AAV28653 Human Cytoplasmic
2	3573	98.8	808	20	AAV28652 Human Cytoplasmic
3	3543.5	98.0	807	20	AAW89247 Human PTP04. Homo
4	2971.5	82.2	799	22	AAW78623 Human tyrosine pho
5	2356.5	65.2	802	20	AAV28654 Murine Cytoplasmic
6	1158	32.0	780	22	AAW51205 Human PTP-PEST Gen
7	1157	32.0	780	21	AAV67252 Human protein tyro
8	1150	31.8	780	22	AAW51201 Human PTP-PEST mut
9	1148	31.8	780	22	AAW51202 Human PTP-PEST mut
10	1147	31.7	780	21	AAV67253 Human mutant PTP-P
11	1127.5	31.2	775	21	AAV67250 Mouse protein tyro

12	1117.5	30.9	775	21	AAV67251 Mutant mouse PTP-P
13	1058	29.3	331	21	AAW58898 Breast and ovarian
14	933	25.8	312	22	AAW78281 Human PTP-PEST. H
15	921.5	25.5	253	22	AAW59383 Human protein tyro
16	748	20.7	433	19	AAW49906 Rat protein tyrosi
17	736.5	20.4	453	18	AAW37254 Novel non-receptor
18	708	19.6	458	19	AAW49908 Human brain derive
19	509	14.1	853	19	AAW85031 Protein tyrosine p
20	508	14.1	1337	16	AAW85203 huDPP-1. Homo sap
21	502	13.9	1149	22	AAW41048 Human polypeptide
22	496	13.7	1440	20	AAV29591 Human protein phos
23	493	13.6	274	22	AAW59385 Protein tyrosine p
24	492.5	13.6	341	22	AAW78283 PTP1. Unidentifie
25	492.5	13.6	1439	15	AAW63632 Human receptor-tyr
26	487	13.5	595	17	AAW99312 Human SH-PTP1 (pro
27	486.5	13.5	593	15	AAW52991 Human protein-tyro
28	485	13.4	595	22	AAW59216 SHP-1 protein. Un
29	485	13.4	595	22	AAW59232 SHP-1 activated tr
30	485	13.4	595	22	AAW59233 SHP-1 activated tr
31	485	13.4	595	22	AAW59238 SHP-1 D59A mutant
32	485	13.4	595	22	SHP-1 E74A mutant
33	485	13.4	597	21	Human Src homolog
34	485	13.4	621	21	Human HCP phosphod
35	485	13.4	824	21	Human HCP/GST fusi
36	485	13.4	843	19	Green fluorescent
37	477	13.2	595	22	SHP-1 activated do
38	477	13.2	595	22	SHP-1 activated do
39	477	13.2	595	22	SHP-1 activated qu
40	477	13.2	595	22	SHP-1 activated qu
41	477	13.2	623	22	Human protein sequ
42	477	13.2	1452	22	Human protein SEQ
43	477	13.2	1455	22	Human protein SEQ
44	476	13.2	1445	15	Human RPTP-gamma.
45	475.5	13.2	1237	19	Chicken protein ty

ALIGNMENTS

RESULT 1	
AAV28653	
ID	AAV28653 standard; Protein; 692 AA.
XX	
AC	AAV28653;
XX	
DT	01-OCT-1999 (first entry)
XX	
DE	Human Cytoplasmic phosphatase, Lyp2 protein.
XX	
KW	Lymphoid Protein Tyrosine Phosphatase; Lyp protein; immunosuppressant;
KW	intracellular tyrosine phosphatase; PTPase; fetal liver; transplamt;
KW	resting lymphoid cell; protein tyrosine kinase; PTKs; lymphocyte;
KW	T cell antigen receptor signalling; cytokine receptor signalling;
KW	autoimmune disease; intronic sequence; alternative mRNA splicing;
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Domain
FT	Location/Qualifiers
FT	27..289
FT	/label= PTPase_domain
FT	/note= "Single catalytic protein tyrosine phosphatase
FT	domain"
FT	469..472
FT	/label= NXXY_motif
FT	/note= "Unique sequence recognised by phosphotyrosine
FT	binding (PTB) domain"
FT	615..623
FT	/label= SH3_binding_site
FT	/note= "Proline rich sequence"
XX	
PN	WO9936548-A1.
XX	
PD	22-JUL-1999.

XX PF 18-JAN-1999; 99WO-CA00038.
XX PR 16-JAN-1998; 98CA-2220853.
XX PA (HSCR-) HSC RES & DEV LP.
XX PI Roifman CM;
XX WPI: 1999-444404/37.
XX N-PSDB; AAX90696.
XX New nucleic acid encoding intracellular tyrosine phosphatase and related proteins, used to modulate signalling through T cells, particularly as immunosuppressant
XX Claim 4b; Page 55; 105pp; English.
XX The present protein sequence is that of the cytoplasmic, Lymphoid Protein Tyrosine Phosphatase, Lyp2 protein that has a single catalytic domain. The non-catalytic portion of the phosphatase contains unique sequences, including a single PEST sequence rich in Pro, Glu or Asp, Ser and Thr. It is expressed significantly in fetal liver and in resting lymphoid cells. Lyp2 is an isoform of the Lyp1 gene that arises by alternative splicing of the mRNA. The intronic sequence of Lyp1 encodes for the C-terminal 7 aminoacids and part of the 3'/untranslated region of Lyp2. Lyp2 shares sequence identity with the murine phosphatase 270PEP. Lyp proteins are important for regulation of T cell antigen and cytokine receptor signalling and for early and late stages of T cell differentiation. Lyp2 has immunosuppressive activity. Compounds that increase expression of Lyp protein can be used as immunosuppressive agents to reduce or prevent T cell activation or proliferation, to control thymocyte differentiation, to treat autoimmune diseases and transplant situations.
XX SQ Sequence 692 AA;
Query Match 100.0%; Score 3615; DB 20; Length 692;
Best Local Similarity 100.0%; Pred. No. 8.2e-300;
Matches 692; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDQREILQKFLDEAQSQKIKTEFANFELKLRQSTKYKADKYPTTVAENAKNIKNRY 60
DB 1 mdgreilqkfldeaqsqskikteefaneflklrqstkykadkypttvaenakniknry 60
QY 61 KDILPYDSRVLSLTSDESSYINANFTKGVYGPAYATOGPLSTLLDFWRMIWEY 120
DB 61 kdilpydsrvlseltsdesssyinanftkgvygpayatogplstlldfwrmiwey 120
QY 121 SVLIIVNACMEYEMGKKKCERYWAEFGEMQLERGFPSVSCAEKRKSDYIIRTLKVKFNS 180
DB 121 svliivnacmeyemgkkkcerywaepegemqlerfgpsvsceaeekrksdyiirtlkvkfns 180
QY 181 ETRTIYOFHYKNPDHPDVPSSIDPILEIWDVRCYQEDDSVPICHCISACGGRGTGICAI 240
DB 181 etrtiyofhyknpdhpdvpsidpileiwdvrcyqeddsvpichcisaagcgrgtgical 240
QY 241 VDYTWMLLKDGIIPENFSVFSLSIREMTQPSLVQTOQEYELVYNVLELFKQMDVIRD 300
DB 241 vdytwmlldkgilpenfsvflsiremtqrpslvqtqgeyelvynvlelfkrqmdvird 300
QY 301 KHSGETSQAKHCIPKNTLIQADSYSNLPKSTTKAAKMMNQORTKMEIKESSSDFRFTS 360
DB 301 khsgetsqakhcipekntliqadsysnlpkstktaakmmnqortkmeikesssdftrfts 360
QY 361 EISAKEELVHPAKSSYFDFLELUNYFDKNADDTMTKQYKAFPIVGEPIQKHQSILDLGS 420
DB 361 eisakeelvlhpakssyfdflenynfydknadtdmtkqykafpivgepiqkhqsidlgs 420
QY 421 LLFEGCSNSKPVNAGRYFNYSKVPITRTKTPPELIQOQRETKEVDKSNFYLESOPHDS 480
DB 421 llfegcsnspvnagryfnyskvpitrtktppeliqoqretkevdknsfylesqphds 480

QY 481 CFVEMOAKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGYSYIPLVENPYFSS 540
DB 481 cfvemgakvmhvssaelnyslpydskhqirnasnvkhhdsalgvysyiplvenpyfss 540
QY 541 WPPSGTSSKMSIDLPEKQDGTVPFSSLLPTSSTLSFYSYNSHSSLSLNSPTNMISSLLNQE 600
DB 541 wppsgtsskmsidlpekqdggtvfpssllptsstlsfysynshsslslnsptnmissllnqe 600
QY 601 SAVLATAPRIDDEIPPLPVRTPESEFIVVEAGEESPVPKSLSSAVKVKICTSLEWGGT 660
DB 601 saviatapiddeipplpvrtpesefivveageetspnvpkslssavkvkigtslewggt 660
QY 661 SBPKKFDSDSVILRPSKVKLRSPKSGKNFSLW 692
DB 661 sepkkfdsdsvilrpskvsklrspksgknfswl 692
RESULT 2
AAY28652
ID AAY28652 standard; Protein; 808 AA.
XX
AC AAY28652;
XX 01-OCT-1999 (first entry)
DE Human Cytoplasmic phosphatase, Lyp1 protein.
XX Lymphoid Protein Tyrosine Phosphatase; Lyp protein; lymphoid cell;
KW Intracellular tyrosine phosphatase; PTPase; lymphocyte; thymocyte;
KW T cell; B cell; protein tyrosine kinase; PTKs; immunosuppressant;
KW T cell antigen receptor signalling; cytokine receptor signalling;
KW autoimmune disease; transplant.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 27..289 /label= PTPase_domain
FT /note= "Single catalytic protein tyrosine phosphatase domain"
FT Domain 469..472 /label= NXXY_motif
FT /note= "Unique sequence recognised by phosphotyrosine binding (PTB) domain"
FT Binding-site 615..623 /label= SH3_binding_site
FT /note= "Proline rich sequence"
FT Binding-site 694..701 /label= SH3_binding_site
FT /note= "SH3 binding site"
FT Region 702..736 /label= PEST_sequence
FT /note= "Sequence rich in Pro, Glu or Asp, Ser and Thr"
FT Region 741..745 /label= PEST_sequence
FT /note= "Consensus sequence recognised by p34cdc2 kinase"
FT Binding-site 768..772 /label= SH3_binding_site
FT /note= "Proline rich sequence"
FT Binding-site 796..804 /label= SH3_binding_site
FT /note= "Proline rich sequence"
XX WO9936548-A1.
XX PN
XX
PD 22-JUL-1999.
XX
PF 18-JAN-1999; 99WO-CA00038.
XX PR 16-JAN-1998; 98CA-2220853.
XX
XX (HSCR-) HSC RES & DEV LP.

PI	Roifman CM;	
DR	WPI; 1999-444404/37.	
DR	N-PSDB; AAX90695.	
XX		
PT	New nucleic acid encoding intracellular tyrosine phosphatase and	
PT	related proteins, used to modulate signalling through T cells,	
PT	particularly as immunosuppressant	
XX		
PS	Claim 4a; Page 53; 105pp; English.	
XX		
CC	The present protein sequence is that of the cytoplasmic, lymphoid	
CC	Protein Tyrosine Phosphatase Lyp1, that has a single catalytic domain.	
CC	The non-catalytic portion of the phosphatase contains unique sequences,	
CC	including a single PEST sequence rich in Pro, Glu or Asp, Ser and Thr.	
CC	Lyp1 is expressed in the lymphoid cells, particularly in thymocytes,	
CC	mature B and T cells and expression is increased upon activation.	
CC	It is involved in lymphocyte growth and development and is	
CC	phosphorylated in a cell cycle dependent manner. Lyp1 shares about 70%	
CC	sequence identity with the murine phosphatase Z70PEP. Lyp proteins are	
CC	important for regulation of T cell antigen and cytokine receptor	
CC	signalling and for early and late stages of T cell differentiation. Lyp1	
CC	has immunosuppressive activity. Compounds that increase expression of	
CC	Lyp protein can be used as immunosuppressive agents to reduce or prevent	
CC	T cell activation or proliferation, to control thymocyte differentiation	
CC	and to treat autoimmune diseases and transplant situations.	
XX		
SQ	Sequence 808 AA;	
Query Match 98.8%; Score 3573; DB 20; Length 808;		
Best Local Similarity 100.0%; Pred. No. 4.1e-296;		
Matches 685; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MDQREILQKFLDQAQSKTKKEFANFLKLRQSTKYKADKYPTTVAENAKNIKNRY	60
Db	1 mddgriqlkfdeaqsaktkkeefaneflklrgstkykadtpttvaenakniknry	60
QY	61 KDILPYDSRVLSLITSDSSYINANFTKGVYGPAYATGQPLSTLLDFWMLWEY	120
Db	61 kdilpydsrvslitsdssyinanftkgvygpayatgqplstlldfwmliwey	120
QY	121 SVLIIVMACMEYEMGKKCBRYWAPEGMOLFPGFSVSCAEKRSYDIIRTLKVKFNS	180
Db	121 svliivmacmeyemgkkcbrywaepemqlfpgfsvscaeakrksdyiirtlkvkfns	180
QY	181 ETRTIQFHYKNPDDHVPSSIDPILLEIWDVRCYQEDDSVPICHCAGCGRTGVICAI	240
Db	181 etrtiyqfhyknpdhdvpsidpilleliwdvrcyqeddsvpicchcagcgrtgvicai	240
QY	241 VDYTWMLLKDGIIIPENFSVPSLIREMRTQPSLVQTEQYELVYNVAVLEFKRQMDVIRD	300
Db	241 vdytwmlldgiiipenfsvpsliremrtqpslvqtqegylvynavlelfrqmdvird	300
QY	301 KHSGTESQAKHCIPKHNHTLQADSYSPNLPKSTTKAAKMNQORTKMEIKESSSDFPRTS	360
Db	301 khsgetesqakhcipeknhhtlqadyspnlpksttkaaakmnqortkmeikesssdfirts	360
QY	361 EISAKELVLPKAKSTSFDFLEBNYSFDKNADTTMKWQTAPPIVGEPIQKHQSLDLGS	420
Db	361 eisakeelvlpkaksstsfdflebnysfdknadttnkwqtkafpivgepiqkhsldlgs	420
QY	421 LLFEGCSNKPVNAAGRYFNASKYPIPTKSTPPELLIQORETKVDKSNFYSLESQPHDS	480
Db	421 llfegcsnkpvnagryfnaskypiptkstpfellqgretkvdksenfyslesqphds	480
QY	481 CFVEMOAKQVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVNPYFSS	540
Db	481 cfvemaqkvnmhvssaenylslypdydkhqirnasnvkhhdsalgvysyiplvnpfyfss	540
QY	541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNNSHSLSLNSPTNISLLNQE	600
Db	541 wppsgtsskmsldlpekqdgvtvpssllptstslfsyynshsslsinsptnissllnqe	600

QY	601 SAVLATAPRIDDEIPPLPVRTPESTFVVEEAGEFSNPVPSLSSAVKVKIGTSLEWGGT	660
Db	601 savlatapriddei:pplpvrtpesfivveeagefsvpnpkslssavkvkigtslewght	660
QY	661 SEPKKFDSDSVILRPSKSVKLRSPKS	685
Db	661 sepkkfdsdsvilrpsksvklrspks	685
RESULT	3	
AAW89247	AAW89247 standard; Protein; 807 AA.	
XX	AAW89247;	
XX	10-MAR-1999 (first entry)	
XX	Human PTP04.	
XX	PTP04; PTP05; PTP10; SAD; ALP; ALK-7; protein tyrosine phosphatase;	
XX	type I receptor serine/threonine kinase; cancer; leukaemia; lymphoma;	
XX	neurodegenerative disease; neuronal survival; Alzheimer's disease;	
XX	Parkinson's disease; Huntington's disease.	
XX	Homo sapiens.	
XX	WO9849317-A2.	
XX	05-NOV-1998.	
XX	27-APR-1998; 98WO-US08439.	
XX	23-OCT-1997; 97US-0063595.	
XX	28-APR-1997; 97US-0044428.	
XX	20-MAY-1997; 97US-0047222.	
XX	11-JUN-1997; 97US-0049477.	
XX	11-JUN-1997; 97US-0049756.	
XX	18-JUN-1997; 97US-0049914.	
XX	(SUGE-) SUGEN INC.	
XX	App H, Clary D, Courtneidge SA, Hui TH, Jallal B;	
XX	Markby D, Onrust S, Peles E, Plowman GD;	
XX	WPI; 1999-009434/01.	
XX	N-PSDB; AAV81742.	
XX	New nucleic acid encoding specific protein tyrosine phosphatases -	
XX	useful for identifying specific modulators for treatment and	
XX	prevention of cancer and neurodegenerative disease	
XX	Claim 2; Page 151-153; 193pp; English.	
XX	The present invention describes isolated, enriched or purified nucleic	
XX	acids encoding PTP04, SAD, PTP05, PTP10, ALP and ALK-7 proteins. The	
XX	present sequence represents human PTP04. The above proteins, other than	
XX	ALK-7, are protein tyrosine phosphatases (PTPs) and are used to identify	
XX	substances that modulate their activity (i.e. agonists and antagonists,	
XX	including NBP) in vivo or in vitro. These substances are used to treat	
XX	or prevent diseases associated with abnormal signal transduction	
XX	pathways that involve the proteins, particularly cancer (e.g. leukaemia	
XX	and lymphoma), while modulators of ALK-7 (which is a type I receptor	
XX	serine/threonine kinase) are used to promote neuronal survival,	
XX	particularly for treating Alzheimer's, Parkinson's or Huntington's	
XX	diseases. Nucleic acid fragments of the polynucleotides encoding the	
XX	proteins can be used as probes to identify and clone related sequences;	
XX	to detect protein-encoded RNA; to generate transgenic animals and in	
XX	gene therapy (optionally after mutation). Ab are used to determine the	
XX	proteins.	
SQ	Sequence 807 AA;	

Query Match 98.0%; Score 3543.5; DB 20; Length 807;
Best Local Similarity 99.4%; Pred. No. 1.4e-293;
Matches 681; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDQREILQKFLDEAOSKKTKEEFANEFUKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60
DB 1 mdqreilqkfldeagskktkeefanefuklrqstkykadttypttvaekpknikknry 60

QY 61 KDILPYDYSRVLSITDESSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
DB 61 kdilpydysrvlsitdsdssyinanfikgygpkayiatqgplstclldfwmiew 120

QY 121 SVLIIVMACMEYEMGKKCERYWAEPMQLEFGPFSVCEAEKRSKSDYIIRTLVKVFN 180
DB 121 svliivmacmeyemgkkcerywaepemqlefgpfsvceaeakrksdyiirtlvkvn 180

QY 181 ETRTIYQFHYKNWPDHVPSSIDPILLEIWDVRCYQEDDSVPICHSCAGCGRTGVICAI 240
DB 181 etrtiyqfhyknwphdvpssidpilleiwdvrcyqeddsvpicihscagcgrtgvicai 240

QY 241 VDTYTMLLKDGIIIPENFVSILIREMRTQPSILVOTQOYELVYNVAVLELFRQMDVIRD 300
DB 241 -dytmllkdgiiipenfvsiliremrtqpsilvotqoeyelvynvavlelfrqmdvird 299

QY 301 KHSGETSQAKHCIPKHNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFRTS 360
DB 300 khsgetsqakhcipeknhtloadsyspnlpksttkaakmmnqortkmeikesssdfrts 359

QY 361 EISAKELVHLPKASTSDFLELNYSPDKNADTTMKWOTKAPPIVGEPLQKHSIDLGS 420
DB 360 eisakeelvhlpkastsdflelnyspdknadttmkwtkafpivgeplqkhsidlgs 419

QY 421 LLFEGCSNKPVNAAGRYNSKVPITRTKSTPELIIQORETKEVDSKFNFSYLESOPHDS 480
DB 420 llfegcsnkpvnagrynskvpitrtkstpfelliqoretkevdskenfylesophs 479

QY 481 CFVEMOQAKVMHVSSAELNYSILPYDSKHQIRNASNVKHHDSALGYSYIPLVENPYFSS 540
DB 480 cfvemqakvmhvssaelnysilpydskhqirnasnvkhhdsaalgyysiplenpyfss 539

QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNNHSSLSLNSPTNSSLNQE 600
DB 540 wppsgtsskmsldlpekqdggtvfpsllptstslfsynnshslslnsptnsslnde 599

QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWEGGT 660
DB 600 savlatapriddeipplpvrtpesfivveeagefspnvpkslssavkvkigtslweggt 659

QY 661 SEPKKFDSDSVILRPKSKVLRSKPS 685
DB 660 sepkkfdsdsvilrpsksvklrpsks 684

RESULT 4
AAG78623
ID AAG78623 standard; Protein; 799 AA.
AC AAG78623;
XX
XX
XX
DT 04-DEC-2001 (first entry)
XX
DE Human tyrosine phosphatase.
XX
XX Human; tyrosine phosphatase; hTTP.
OS Homo sapiens.
XX
XX CN1302899-A.
XX
PD 11-JUL-2001.
XX
PF 29-OCT-1999; 99CN-0119935.

XX 29-OCT-1999; 99CN-0119935.
XX (NANF-) NANFANG RES CENT STATE HUMAN GENE GROUP.
XX
XX Ren S, Wu T, Qian B;
DR WPI; 2001-550591/62.
DR N-PSDB; AAH79342.
XX Human protein tyrosine phosphatase and its coding sequence -
PT Claim 4; Page 12-13(Disclosure); 29pp; Chinese.
PS
XX The present invention provides the protein and coding sequences of human
CC tyrosine phosphatase hTTP. The protein is expressed in human normal
CC suprarenal tissue. The present sequence is the protein of the invention.
XX
SQ Sequence 799 AA;

Query Match 82.2%; Score 2971.5; DB 22; Length 799;
Best Local Similarity 87.3%; Pred. No. 1e-244;
Matches 598; Conservative 17; Mismatches 61; Indels 9; Gaps 9;

QY 1 MDQREILQKFLDEAOSKKTKEEFANEFUKLRQSTKYKADKTYPTTVAENAKNIKKNRY 60
DB 1 mdqreilqkfldeagskktkeefanefuklrqstkykadttypttvaekpknikknry 60

QY 61 KDILPYDYSRVLSITDESSYINANFIKGVYGPAYIATQGPLSTLLDFWMIWEY 120
DB 61 kdilpydysrvlsitdsdssyinanfikgygpeayiatqgplstlldfwmiew 120

QY 121 SVLIIVMACMEYEMGKKCERYWAEPMQLEFGPFSVCEAEKRSKSDYIIRTLVKVFN 180
DB 121 svliivmacmeyemgkkcerywaepemqlefgpfsvceaeakrksdyiirtlvkvn 180

QY 181 ETRTIYQFHYKNWPDHVPSSIDPILLEIWDVRCYQEDDSVPICHSCAGCGRTGVICAI 240
DB 181 etrtiyqfhyknwphdvpssidpilleiwdvrcyqeddsvpicihscagcgrtgvicai 240

QY 241 VDTYTMLLKDGIIIPENFVSILIREMRTQPSILVOTQOYELVYNVAVLELFRQMDVIRD 300
DB 241 -dytmllkdgiiipenfvsiliremrtqpsilvotqoeyelvynvavlelfrqmdvird 299

QY 301 KHSGETSQAKHCIPKHNHTLOADSYSPNLPKSTTTAAKMMNQORTKMEIKESSSDFRTS 360
DB 300 khsgetsqakhcipeknhtloadsyspnlpksttkaakmmnqortkmeikesssdfrts 359

QY 361 EISAKELVHLPKASTSDFLELNYSPDKNADTTMKWOTKAPPIVGEPLQKHSIDLGS 420
DB 360 eisakeelvhlpkastsdflelnyspdknadttmkwtkafpivgeplqkhsidlgs 416

QY 421 LLFEGCSNKPVNAAGRYNSKVPITRTKSTPELIIQORETKEVDSKFNFSYLESOPHDS 480
DB 417 lcl-gmsnlnv-csrkifnprcqp-dqinsfeliqq-enggdskkttlf-esqphs 471

QY 481 CFVEMOQAKVMHVSSAELNYSILPYDSKHQIRNASNVKHHDSALGYSYIPLVENPYFSS 540
DB 472 cfvemqakvmhvssaelnysilpydskhqirnasnvkhhdsaalgyysiplenpyfss 531

QY 541 WPPSGTSSKMSLDLPEKQDGTVPSSLLPTSTSLFSYNNHSSLSLNSPTNSSLNQE 600
DB 532 wppsgtsskmsldlpekqdggtvfpsllptstslfsynnshslslnsptnsslrieqe 591

QY 601 SAVLATAPRIDDEIPPLPVRTPEFIVVEEAGEFSPNPKSLSSAVKVKIGTSLWEGGT 660
DB 592 savlatapriddeipplpvrtpesfivveeagefspnvpkslssavkvkigtslweggt 651

QY 661 SEPKKFDSDSVILRPKSKVLRSKPS 685
DB 652 sepkkfdsdsvilrpsksvklrpsks 676

RESULT 5
 AAY28654
 ID AAY28654 standard; Protein; 802 AA.
 XX
 AC AAY28654;
 XX
 DT 01-OCT-1999 (first entry)
 XX
 DE Murine Cytoplasmic phosphatase, Z70PEP protein.
 XX
 XX Lymphoid Protein Tyrosine Phosphatase; Lyp protein; lymphoid cell;
 KW intracellular tyrosine phosphatase; PTPase; lymphocyte; murine;
 KW protein tyrosine kinase; PTKs; immunosuppressant; PEST sequence;
 KW T cell antigen receptor signalling; autoimmune disease; transplant;
 KW cytokine receptor signalling.
 XX
 OS Mus sp.
 FH
 FH Key Location/Qualifiers
 FT Domain /label= PTPase_domain
 FT /note= "Catalytic protein tyrosine phosphatase domain"
 FT Binding-site 613..621
 FT /label= SH3_binding_site
 FT /note= "Proline rich sequence"
 FT Binding-site 689..695
 FT /label= SH3_binding_site
 FT /note= "Proline rich sequence"
 FT Binding-site 790..798
 FT /label= SH3_binding_site
 FT /note= "Proline rich sequence"
 XX
 WO9936548-AL.
 PN
 XX
 XX 22-JUL-1999.
 XX
 XX 18-JAN-1999; 99WO-CA00038.
 XX
 PR 16-JAN-1998; 98CA-2220853.
 XX
 XX (HSCR-) HSC RES & DEV LP.
 PA
 PI Roifman CM;
 XX
 XX WPI; 1999-444404/37.
 XX
 XX New nucleic acid encoding intracellular tyrosine phosphatase and
 PT related proteins, used to modulate signalling through T cells,
 PT particularly as immunosuppressant
 XX
 PS Disclosure; Page 63-64; 105pp; English.
 XX
 CC The present protein sequence is that of the murine phosphatase, Z70PEP
 CC that has a single catalytic domain. The non-catalytic portion of the
 CC phosphatase contains unique sequences, including five PEST sequences
 CC rich in Pro, Glu or Asp, Ser and Thr. Z70PEP shares about 70% sequence
 CC identity with the human cytoplasmic phosphatase Lyp1. Lyp proteins are
 CC important for regulation of T cell antigen and cytokine receptor
 CC signalling and for early and late stages of T cell differentiation.
 CC Z70PEP has immunosuppressive activity. Compounds that increase
 CC expression of Lyp protein can be used as immunosuppressive agents to
 CC reduce or prevent T cell activation or proliferation, to control
 CC thymocyte differentiation, to treat autoimmune diseases and transplant
 CC situations.
 XX
 XX Sequence 802 AA;
 SQ
 Query Match 65.2%; Score 2356.5; DB 20; Length 802;
 Best Local Similarity 68.5%; Pred. No. 3.9e-192;
 Matches 474; Conservative 66; Mismatches 139; Indels 13; Gaps 6;

QY 1 MDQREILQKFLDEAQAQKKITKEEFANEFKLKRSQSTKYKADKTYPTTVAENAKNIKNRY 60
 DB 1 mdqreilqllkeaqkklaseefaseflklkrqstkykadypttvaqrpnknkny 60
 QY 61 KDILPYDYSRVLSLITSDSSSYINANFIKGYGPKAYIATOGPLSTTLLDFWRMIWEY 120
 DB 61 kdilpydhsylslitsdsssyinasfikgygpkayiatdgpsttlldfwrmiwey 120
 QY 121 SVLIIVMACMEYEMGKKCERYWAEPCGEMOLEFGPPSVSCAEKRSYDIIRTLKVFNS 180
 DB 121 rliivimacmefemgkckerywaegetdqlf9pfsisceaeakksdyirtlkakfnn 180
 QY 181 ETRTIYQFHYKNWPDHDPSSIDPILLEIWDVRCYQEDDSVPICIHGACGRTGVICAI 240
 DB 181 etriiyqfhyknwpdhdpssidpilqlwdmrcygeddcvpcihcsagcgrtgvisa- 239
 QY 241 VDYTWMLLKDGIIIPENFVSFLIREMRTQPSLVQTOEQEYELVYNVLELFKROMDIVRD 300
 DB 240 vdytwmlldgiiipknfsvnlqemrtqpslvqtgeqyelvysavlelfkthmdvisd 299
 QY 301 KHSGETSQAKHICPEKNHTLQADSYSNLPKSTTKAAKMMNQ---ORTKMEIKESSSDF 357
 DB 300 nhlgreiqacsipegsltveadscpldlpknamrdvktndhskqgaeaestggsslg 359
 QY 358 RTSEISAKELVLHPAKSSTSFDFLELNYSFKNADTTMKWTKAPPIVGEPLQKHQSLD 417
 DB 360 rtstmaeeelvlhsaksspsfncleincgcnnkavitrngqaraspvgeplqkyqsl 419
 QY 418 LGSLLFEGCSNRPVNAAGRYFNKVPITRTKSTPELIIQORETKEVDSENFYSLESQP 477
 DB 420 fgmllfgscpsalpdintadryhnskgyvkrktatpfelldqgrktndlavdgdfsclesql 479
 QY 478 HDSCFV-EMQAKVMHVSSAELNYSLPYDSKHQIRNASNVKHHDSALGVYSYIPLVENP 536
 DB 480 hehysirelqvrvahvsseelnyslp-----gacdacvprhspgairvhllytslaedp 534
 QY 537 YFSWPPPSGTSKMSLDLPEKQDGTVPFSSLLPTSTLSFYNSHSSLSLNSPTNTSSL 596
 DB 535 yfssppnsadskmsfdlpekqdgatspgallpassttsffynphdslymntltsfpp 594
 QY 597 LNOESAVLATAPRIDDEIPPPPLVPTPESFIVVEEAGEFSPNVPKSLSSAVKVIKIGTSLE 656
 DB 595 lndetaveapsrrtddeipplpertsfivveeagepsprvtesl--plvvtfgaspe 652
 QY 657 WGGTSPKPKFDDSVILRPKSVKLRGPKSGKN 688
 DB 653 csytsemks-hdsvgftpsknvklrpsksdrh 683
 RESULT 6
 AAM51205
 ID AAM51205 standard; Protein; 780 AA.
 XX
 AC AAM51205;
 XX
 DT 19-DEC-2001 (first entry)
 XX
 DE Human PTP-PEST Genbank Accession Number XP034191.
 XX
 KW PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
 KW dephosphorylation; phosphotyrosine; human; PTP-PEST.
 OS Homo sapiens.
 XX
 XX WO200161031-A2.
 PN
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US05180.
 XX
 PR 14-FEB-2000; 2000US-0181769.
 XX
 PA (CEPT-) CEPTYR INC.

XX Flint AJ, Cool DE;
XX WPI; 2001-570570/64.
XX Screening assays to identify agents that alter protein tyrosine
XX phosphatase (PTP) binding to, and PTP-mediated catalytic
XX dephosphorylation of phosphotyrosine peptide substrates -
XX Example 1; Page -; 79pp; English.
XX The invention relates to identifying agents which alter the interaction
XX between a protein tyrosine phosphatase (PTP) and a tyrosine
XX phosphorylated polypeptide using fluorescence energy signals. The methods
XX are useful for performing screening assay to identify agents that alter
XX PTP binding to and PTP-mediated catalytic dephosphorylation of
XX phosphotyrosine peptide substrates. The present sequence is not given in
XX the specification but is that of human PTP-PEST protein sequence taken
XX from Genbank (Accession Number; XP034191). The present sequence was used
XX to generate mutants D199A (AAM51201) and C231S (AAM51202) as described in
XX the specification.
XX Note: An alternative sequence for human PTP-PEST is given in figure 1 of
XX the specification (AAG78281).
XX Sequence 780 AA;

Query Match 32.0%; Score 1158; DB 22; Length 780;
Best Local Similarity 37.2%; Pred. No. 1e-89;
Matches 291; Conservative 111; Mismatches 237; Indels 144; Gaps 25;
QY 1 MDQREIQFLDEAOSKKT-----KEEFANEFKLKROSTKYKADKTYPTTVAENAKNIK 56
DB 1 meqveillrficrvqamkspdhngednfardfmriristkyrtktyrtatgekeenvk 60
QY 57 KRYKRDILPYDSRVLSLITDESSYNANFIVGVGPKAYIATQGPLSTLLDFWRM 116
DB 61 krykdilpfdhsrvklitktpsgdsdyinanfkgvypkayvatqgplantvdfwr 120
QY 117 IWEYSLLIIVMACMEYEMGKCEKRYWABEGMQLFPGFVSCAEKRSKYIIRTLKV 176
DB 121 iweynvllvmacreferemgkceerywlygedpdtfapfkiscdeqartdyfirtl 180
QY 177 KFNSERTTYQHYKNWPDHVPSSIDPLELIMVRCYQDDSPICHCASAGCGRTGV 236
DB 181 efqnesrlyqhyvnpdhdvpsfidsldmslmrkyqehedvpichcsagcgtrga 240
QY 237 ICAIVDYTWMLLDGIIIPENSVFSLIREMRTQPSLQTOQYELVYNNAVLEFKRQMD 296
DB 241 icai-dytwnllkagkipeefnvnligemrtgrhsavqtkeqyelvhraiaqlfekql 299
QY 297 V-----IRD--KHSGVTSQAKHCIPKNTLQADSYSPNLPKSTTKAAKMMNQORTK 346
DB 300 lyeihgaqkiadgvneintemvssiepek-----dspppkprrts-----clvegda 350
QY 347 MEIKE-----SSSDFRTSEISAKEEL-----VLHPAKSSTSFDFLELNY 387
DB 351 eellqppehpvppiltpspasfptvtvwdndryhpkvplhmvsseqhsadlnnys 410
QY 388 -----FDKNADTMKQWTAFPIVGEPLQKHOSLDLGSLLFEG----- 425
DB 411 kstelpgknestieqdkdklernalsfeikvplqegp-----ksfdgntllnrghaik 466
QY 426 ---C-----SNSKP-----VNAAGRYNSKVPIITRTKS-----TPFELIQO 462
DB 467 aspciadkiskqogelsadlnvgdtsqscvdcsvtknkvsvtppsqsndtpprpdrl 526
QY 463 EVDSKENFYLESPHDSCFVEMQAQVMHVSAAELNY-----SLPYDSKHQIRNASV 516
DB 527 pldekghvtfwsgfngpna-----ipidlsegnssdinyqtrktvsltpspqtqvdpdl 582
QY 517 KHHDSALGVYVIPLENVPPYSSWPPSGTSSKMSLDLPEK-QDGTVPFSSLLPSSSTSL 575

DB 583 dhndts-----plfrtplsftnplhdsds-----dsdernsdgavtqknknstasat 630
QY 576 FSYNSSHSLNSPTNISLLNOESAVLATAPR-----IDDEIPPLPVRTPESFIVVEE 631
DB 631 vsaatstesistrkvlpmislarhniagtthsgaekdvdsedspplpertsfvlase 690
QY 632 AGEFSPNVPKSLSSAVKVKIGTSLWEG-----GTSEPKKFDSDSVILRPSKSVKLSRPSKG 686
DB 691 -----hntp-----vrsewselqserseqkkseg-----litsenekcdhpagg 730
QY 687 KNF 689
DB 731 lhy 733
RESULT 7
AAY67252
ID AAY67252 standard; protein; 780 AA.
XX AC AAY67252;
XX DT 05-APR-2000 (first entry)
XX DE Human protein tyrosine phosphatase (PTP) PEST amino acid sequence.
XX KW Protein tyrosine phosphatase; PTP; PEST; paxillin binding; adhesion;
XX KW cell migration; division; cytosolic; antiinflammatory; angiogenesis;
XX KW cancer; enzyme substrate identification; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 333..340
FT /label= Pro_1
FT /note= "Proline rich domain 1"
FT Domain 356..365
FT /label= Pro_2
FT /note= "Proline rich domain 2"
FT Domain 519..528
FT /label= Pro_3
FT /note= "Proline rich domain 3"
FT Domain 674..680
FT /label= Pro_4
FT /note= "Proline rich domain 4"
FT Domain 769..776
FT /label= Pro_5
FT /note= "Proline rich domain 5"
XX WO9961467-A2.
XX PN 02-DEC-1999.
XX PD 21-MAY-1999; 99WO-CA00461.
XX PF 21-MAY-1998; 98CA-2238654.
XX PR 11-DEC-1998; 98US-0111993.
XX PA (UYMC-) UNIV MCGILL.
XX PI Tremblay ML, Cote J, Angers-Lousteau A, Charest A;
XX WPI; 2000-097104/08.
XX Novel therapeutic agents for treating diseases which are related to
XX cell proliferation, migration, inflammation and angiogenesis especially
XX cancer -
XX Claim 4; Fig 24; 91pp; English.
XX This is the amino acid sequence of human protein tyrosine phosphatase
XX (PTP) PEST. PTP-PEST is a soluble PTP that is ubiquitously expressed
XX throughout embryonic development and in murine adult tissues. The
XX N-terminal portion of the enzyme encodes for the catalytic domain, while


```

691      | |      | |      | | | | : | | : | |
      http-----vrsewslqsrsekksq---litsenekcdhpqg 730
Db

```

DE
Human PTP-PEST mutant D199A.
XX

XX	dephosphorylation; phosphotyrosine; human; FIF-FEST; mutant.
XX	
OS	Homo sapiens.
OS	Synthetic.

Key	Location/Qualifiers
Misc-difference	199
FT	/note= "Wild-type Asp substituted by Ala"
PN	WO200161031-A2.
XX	23-AUG-2001.
XX	13-FEB-2001; 2001WO-US05180.
XX	14-FEB-2000; 2000US-0181769.
PR	(CEPT-) CEPTYR INC.
PA	Flint AJ, Cool DE;
XX	WPI; 2001-570570/64.
DR	Screening assays to identify agents that alter protein tyrosine phosphatase (PTP) binding to, and PTP-mediated catalytic dephosphorylation of phosphotyrosine peptide substrates -
PT	Example 1; Page -: 79pp; English.
XX	The invention relates to identifying agents which alter the interaction between a protein tyrosine phosphatase (PTP) and a tyrosine phosphorylated polypeptide using fluorescence energy signals. The method are useful for performing screening assay to identify agents that alter PTP binding to and PTP-mediated catalytic dephosphorylation of phosphotyrosine peptide substrates. The present sequence is that of a human PTP-PEST D199A mutant.
CC	Note: The present sequence is not given in the specification but is derived from the wildtype human PTP-PEST sequence (AAM51205) from Genebank (Accession Number: XP034191). An alternative sequence for human PTP-PEST is given in figure 1 of the specification (AAG78281).
CC	Sequence 780 AA;
XX	Query Match 31.8%; Score 1150; DB 22; Length 780;
XX	Best Local Similarity 37.0%; Pred. No. 5e-89;
XX	Matches 290; Conservative 111; Mismatches 238; Indels 144; Gaps
QY	1 MDQREILQKFLDQAQSKKIT- ---KEEFANEFLKLRQSTKYKADKTYPTTVAENAKNIK 56
Db	1 meqveilrkfiqrqamkspdhngednfmrlrlstkyrtekiyptatgeenvk 60
QY	57 KNRYKDILPYDYSRVLSLTSDEDSYINANFIKGYGPKAVIATQGPLSTLLDFWRM 116

Db 61 knrykdilpfdhsrvklltlktspqdsdyinanfikvgypkayvatgqplantvidfwm 120
QY 117 IWEYSVLIIVMACMEYEMGKKKCYWAEPMQEMQLEFGPFSVSCBAEKRSYIIIRTLKV 176
Db 121 lweynvliivmacrefemgrkccerywplygedpitfapfkiscdeqartdyfirtlil 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILILEIWDVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 efqnesrrlyqfhyvnpahdvpssfdslmdislmrkyqehedvpicihcsagcgrtga 240
QY 237 ICAIVDYTWMLKDGIIIPENFSVSLIREMRTQPSLQTOQOYELVYNAVLELFRQMD 296
Db 241 lcai-dytnllkagkipeefnvnfligemrtqrsavqtkeqyelvhratlaqlfekqlq 299
QY 297 V-----IRD--KHSGTESQAKHCIPKNTHTLQADSYSPNLPKSTTKAAKMMNQORTK 346
Db 300 lyeihgaqkiadgvneintemvssiepek-----dspppkpptrs-----clvegda 350
QY 347 MEIKE-----SSSDFRTSEISAKEL-----VLHPAKSSTSFDFLELNY 387
Db 351 eellqppephvpypalltpspasfptvtvwdndryhpkpvlhmvsseqhsadlnrnys 410
QY 388 -----FDKNADTTMKWTKAFPIVGEPLQKHQSLDLSLLFEG----- 425
Db 411 kstelpgknestieqldkklernlsfeikkvplqegp-----ksfdgntllnrghaikls 466
QY 426 ---C---SNSKP-----VNAAGRYFNKSVPIRTKTS-----TPFELIQORETK----- 462
Db 467 aspiadkiskpqelssdinvgdtsgnscvdcsvtqskvsvtppesqnsdtprrdrl 526
QY 463 EVDSENFSYLSQPHSDSCFVEMQAKVMHVSSEALNY-----SLPYDSKHQIRNASV 516
Db 527 pldkghvtwsthgpena---ipldiseqnsdinyqtrkvtsltpsttqvtpdlv 582
QY 517 KHDSSALGVSYIPLVNPENYSSPPSGTSSKMSLDLPEK-QDGTVPFSSLLPTSSSL 575
Db 583 dhndts-----plfrclftnplhsds-----dsdernsdgvatqknkntasat 630
QY 576 FSYNHSLSLSNPSNTNISLLNQESAVLATAPR-----IDDEIPPLPVRTPESTFVBE 631
Db 631 vsaatstesistrkvlpmisrhiarhiagtthsgaekdvdsedspplpertsfvlase 690
QY 632 AGEFSNPVKSLSAVKVKIGTSLEWG-----GTSEPKFDDSVILRPSKVKLSRPSKG 686
Db 691 -----hntp-----vrsewselqserseqkkseg---litsenekcdhpagg 730
QY 687 KNF 689
Db 731 lhy 733

RESULT 9
AAM51202
ID AAM51202 standard; Protein; 780 AA.
XX
AC AAM51202;
XX
DT 19-DEC-2001 (first entry)
XX
DE Human PTP-PEST mutant C231S.
XX
KW PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
KW dephosphorylation; phosphotyrosine; human; PTP-PEST; mutain; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 231
FT /note= "Wild-type Cys substituted by Ser"
XX
PN
XX WO200161031-A2.

PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US05180.
XX
PR 14-FEB-2000; 2000US-0181769.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Flint AJ, Cool DE;
XX
DR WPI; 2001-570570/64.
XX
PT Screening assays to identify agents that alter protein tyrosine
PT phosphatase (PTP) binding to, and PTP-mediated catalytic
PT dephosphorylation of phosphotyrosine peptide substrates -
XX
XX Example 1; Page -: 79pp; English.
XX
CC The invention relates to identifying agents which alter the interaction
CC between a protein tyrosine phosphatase (PTP) and a tyrosine
CC phosphorylated polypeptide using fluorescence energy signals. The methods
CC are useful for performing screening assay to identify agents that alter
CC PTP binding to and PTP-mediated catalytic dephosphorylation of
CC phosphotyrosine peptide substrates. The present sequence is that of a
CC human PTP-PEST C231S mutant.
CC Note: The present sequence is not given in the specification but is
CC derived from the wildtype human PTP-PEST sequence (AAM51205) from
CC Genbank (Accession Number; XP034191). An alternative sequence for human
CC PTP-PEST is given in figure 1 of the specification (AAG78281).
XX
SQ Sequence 780 AA;

Query Match 31.8%; Score 1148; DB 22; Length 780;
Best Local Similarity 37.0%; Pred No. 7.4e-89;
Matches 290; Conservative 111; Mismatches 238; Indels 144; Gaps 25;

QY 1 MDQREILQKFLDEAQSCKIT----KEEFANEFKLKROSTKYKADKYPTTVAENAKNIK 56
Db 1 meqveillrkfqrvcamkspdhngednfardfmrllstkyrtekiyptatgekenvk 60
QY 57 KNRYKDILPYDSRVVELSLITSDSSYINANFIKGVGPKAYIATQGPLSTLLDFWRM 116
Db 61 knrykdilpfdhsrvklltlktspqdsdyinanfikvgypkayvatgqplantvidfwm 120
QY 117 IWEYSVLIIVMACMEYEMGKKKCYWAEPMQEMQLEFGPFSVSCBAEKRSYIIIRTLKV 176
Db 121 lweynvliivmacrefemgrkccerywplygedpitfapfkiscdeqartdyfirtlil 180
QY 177 KFNSETRTIYQHYKNWPDHVPSSIDPILILEIWDVRCYQEDDSVPICIHCSAGCGRTGV 236
Db 181 efqnesrrlyqfhyvnpahdvpssfdslmdislmrkyqehedvpicihcsagcgrtga 240
QY 237 ICAIVDYTWMLKDGIIIPENFSVSLIREMRTQPSLQTOQOYELVYNAVLELFRQMD 296
Db 241 lcai-dytnllkagkipeefnvnfligemrtqrsavqtkeqyelvhratlaqlfekqlq 299
QY 297 V-----IRD--KHSGTESQAKHCIPKNTHTLQADSYSPNLPKSTTKAAKMMNQORTK 346
Db 300 lyeihgaqkiadgvneintemvssiepek-----dspppkpptrs-----clvegda 350
QY 347 MEIKE-----SSSDFRTSEISAKEL-----VLHPAKSSTSFDFLELNY 387
Db 351 eellqppephvpypalltpspasfptvtvwdndryhpkpvlhmvsseqhsadlnrnys 410
QY 388 -----FDKNADTTMKWTKAFPIVGEPLQKHQSLDLSLLFEG----- 425
Db 411 kstelpgknestieqldkklernlsfeikkvplqegp-----ksfdgntllnrghaikls 466
QY 426 ---C---SNSKP-----VNAAGRYFNKSVPIRTKTS-----TPFELIQORETK----- 462
Db 467 aspiadkiskpqelssdinvgdtsgnscvdcsvtqskvsvtppesqnsdtprrdrl 526

[illegible]

```
XX PF 08-MAR-2000; 2000WO-US05881.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-611515/58.
XX DR N-PSDB; AAF21801.
XX PR New human breast and ovarian cancer associated gene sequences and the
XX PT polypeptides encoded by these genes, useful in the prevention,
XX PT treatment and diagnosis of cancer, immune disorders, cardiovascular
XX PT disorders and neurological diseases -
XX PS Claim 11; Page 1045-1047; 1299pp; English.
XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX CC associated with breast and ovarian cancer. Included in the invention are
XX CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX CC isolation and characterisation of the DNA and protein sequences of the
XX CC invention. The breast and ovarian cancer associated DNA, protein, agonist
XX CC or antagonist sequences exhibit cytostatic; immunosuppressive;
XX CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX CC antidiabetic; antinflamatory; antiulcer; antiparasitic and cardiant activity. The
XX CC antibacterial; antifungal; antiparasitic and cardiant activity. The
XX CC polynucleotide and protein sequences are used in the diagnosis of cancer,
XX CC particularly breast and ovarian cancer. The nucleic acid sequences,
XX CC proteins, agonists and agonists may also be used in the diagnosis,
XX CC prevention and treatment of immune disorders e.g. Addison's disease,
XX CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX CC arthritis and ulcerative colitis; cardiovascular disorders such as
XX CC myocardial ischaemias; wound healing; neurological diseases such as
XX CC cerebral anoxia and epilepsy; and infectious diseases.
XX SQ Sequence 331 AA;

Query Match 29.3%; Score 1058; DB 21; Length 331;
Best Local Similarity 99.0%; Pred. No. 8.9e-82;
Matches 206; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 478 HDSCFVMOAQKVMYSSAELNYSPLPDSKHOIRNASNVKHHDSALGVYSYIPLVENPY 537
Db 1 hdscfvemaqaqvmhvssaelnysplpdskhqirnasnvkhhdsalgvysyiplvenpy 60

QY 538 FSWPPSGTSSKMSLDLPKQDQGVFPSSLLPTSTSLFSYYNSHSSLSLSNPTNISLL 597
Db 61 fswppsgtsskmsldlpkqdgvtfpssllxptsstslfsyynshdlslnsptnissll 120

QY 598 NQESAVLATAPRIDDEIPPLVPRTPESTIVVEEAGEFSPNPKLSUAVKVKIGTSLBW 657
Db 121 nqesavlatapriddeipplvprrtpestivveeagefspnvpklsuavkvkigtlslew 180

QY 658 GGTSEPKKFDSDVILRPKSKVLRSKPS 685
Db 181 ggtsepkkfdsdvillrpkskvlrskps 208

RESULT 14
AAG78281
ID AAG78281 standard; Protein; 312 AA.
XX AC AAG78281;
XX DT 19-DEC-2001 (first entry)
XX DE Human PTP-PEST.
XX DT
```

```
KW PTP; protein tyrosine phosphatase; tyrosine phosphorylated polypeptide;
KW KW dephosphorylation; phosphotyrosine; human; PTP1B; mouse; fruit fly;
XX yeast.
XX OS Homo sapiens.
XX PN WO200161031-A2.
XX PD 23-AUG-2001.
XX PF 13-FEB-2001; 2001WO-US05180.
XX PR 14-FEB-2000; 2000US-0181769.
XX PA (CEPT-) CEPTYR INC.
XX PI Flint AJ, Cool DE;
XX DR WPI; 2001-570570/64.
XX CC Screening assays to identify agents that alter protein tyrosine
XX CC phosphatase (PTP) binding to, and PTP-mediated catalytic
XX CC dephosphorylation of phosphotyrosine peptide substrates -
XX PS Disclosure; Fig 1; 79pp; English.
XX CC The invention relates to identifying agents which alter the interaction
XX CC between a protein tyrosine phosphatase (PTP) and a tyrosine
XX CC phosphorylated polypeptide using fluorescence energy signals. The methods
XX CC are useful for performing screening assay to identify agents that alter
XX CC PTP binding to and PTP-mediated catalytic dephosphorylation of
XX CC phosphotyrosine peptide substrates. The present sequence is that of a
XX CC catalytic domain of a PTP for comparison with human PTP1B (AAG78262).
XX SQ Sequence 312 AA;

Query Match 25.8%; Score 933; DB 22; Length 312;
Best Local Similarity 54.2%; Pred. No. 3.9e-71;
Matches 182; Conservative 47; Mismatches 57; Indels 50; Gaps 6;

QY 44 YPTTVAENAKNKKRYKDILPYDSRVLSLTSDESSYINANFKGVYGPKAYIATQ 103
Db 2 yptatgekeenkvknrykdilpfdhsrvkltkltpsqdsdyinankfgyvgpkayvatq 61

QY 104 GPLSTTLLDFWRMINEYSVLIIVMACMEYEMGKKCKERYWAPGEMQLFEGPESVCEAE 163
Db 62 gplantvldfwrwmvneynvliivmacrefemgrkckerywplygedpifafkiscde 121

QY 164 KRKSDYIIRTLKVFNSERTTIYQFHYKNWPDHVPSSIDPILELIWDRVCYQEDDSVPI 223
Db 122 qartdyfi-lllefqnesrrlyqfhyvnpdhdvpsfsldmslmrkygehedvpi 179

QY 224 CJHCSAGCGRTGVICAIVDYTWMLKDGIIIPENFVSFLIREMRTORPSLSVQTQOYELV 293
Db 180 cjhsagcgtrtaicai-dytwnllkagkipeefnfnliqemrtgrhsavgtkeqyelv 238

QY 284 YNAVLEFLKQMDVIRDKHSGTESQAKHCIPKKNHTLQADSYSPNLPKSTTKAAK ---- 338
Db 239 hraiaqlfekdq-----lyaidywnl-----lkagkipeef 271

QY 339 ----MMNQORTKMEIKESSSDFRTSEISAKBELVL 370
Db 272 nvfnliqemrtg-----rhsvavtkeqyel 296

RESULT 15
AAB59383
ID AAB59383 standard; Protein; 253 AA.
XX AC AAB59383;
XX DT 21-MAR-2001 (first entry)
```


THIS PAGE BLANK (USPTO)